

**SUOMEN YORKSHIRE-SIKOJEN
HEDELMÄLLISYYSOMINAISUUKSIEN GENOMINEN
ANALYYSI SNP-MARKKEREIDEN AVULLA**

Sini Wallén
Maisterin tutkielma
Helsingin yliopisto
Maataloustieteiden laitos
Kotieläinten jalostustiede
Toukokuu 2012

Tiedekunta/Osasto — Fakultet/Sektion — Faculty		Laitos — Institution — Department	
Maatalous-metsätieteellinen tiedekunta		Maataloustieteiden laitos	
Tekijä — Författare — Author			
Sini Wallén			
Työn nimi — Arbetets titel — Title			
Suomen yorkshire-sikojen hedelmällisyysominaisuuksien genominen analyysi SNP-markkereiden avulla			
Oppiaine — Läroämne — Subject			
Kotieläinten jalostustiede			
Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages	
Maisterin tutkielma	Toukokuu 2012	44	
Tiivistelmä — Referat — Abstract			
<p>Hedelmällisyysominaisuuksien alhaisen periytymisasteen vuoksi näiden ominaisuuksien parantaminen perinteisen jalostusvalinnan avulla on osoittautunut haasteelliseksi, eikä niiden perinnöllinen edistyminen ole täysin vastannut odotuksia myöskään Suomessa. Yksi mahdollisuus hedelmällisyysominaisuuksien perinnöllisen edistymisen nopeuttamiseen on käyttää geenimerkkejä. Assosiaatioanalyysin tuloksena löydetään ne geenimerkit, joilla on tilastollisesti merkitsevä vaikutus tutkittavan ominaisuuden suhteen. Tutkimuksen tavoitteena oli löytää yorkshire-sioilla ne kromosomialueet, jotka ovat yhteydessä sian hedelmällisyysominaisuuksiin. Tarkoituksena oli myös selvittää, ovatko yorkshire-sioista löydetty kromosomialueet yhteydessä samoihin hedelmällisyysominaisuuksiin aiemmin maatiaissioilta löydettyihin alueisiin. Lisäksi tutkimuksessa pyrittiin löytämään tilastollisesti merkitsevien SNP-markkereiden lähellä sijaitsevia geenejä. Tutkimuksessa tarkasteltiin yhdeksää naarashedelmällisyyden ominaisuutta, jotka olivat syntyneiden porsaiden lukumäärä ensimmäisessä ja myöhemmissä pahnueissa, kuolleina syntyneiden porsaiden lukumäärä ensimmäisessä ja myöhemmissä pahnueissa, porsaskuolleisuus syntymästä vieroitukseen ensimmäisessä ja myöhemmissä pahnueissa, emakon ikä ensimmäisellä porsimisella sekä ensimmäinen ja toinen porsimisväli. Ominaisuuksien ja SNP-markkerien välisen assosiaation selvittämiseen käytettiin lineaarista sekamallia, joka sisälsi kiinteän SNP-vaiikutuksen ja satunnaisen polygeenisen vaikutuksen. Assosiaatioanalyysi tehtiin AI-REML-menetelmällä, lisäksi tutkimuksessa käytettiin deregressoituja EBV:itä ja painokertoimia. Tilastollisen merkitsevyyden selvittämisessä käytettiin Bonferronin korjausta. Tutkimuksessa löydettiin 20 sian hedelmällisyysominaisuuksiin vaikuttavaa tilastollisesti merkitsevää (P-arvo $\leq 2,0E-06$) kromosomialuetta. Näistä markkereista yksi (krom. 7) liittyi syntyneiden porsaiden lukumäärään ensimmäisessä pahnueessa (P-arvo = $8,39E-08$), kaksi (krom. 7 ja 1) syntyneiden porsaiden lukumäärään myöhemmissä pahnueissa (P-arvot = $2,77E-07$ ja $1,91E-06$), yksi (krom. 8) ensimmäisessä pahnueessa kuolleina syntyneiden porsaiden lukumäärään (P-arvo = $9,72E-08$), kolme (krom. 13) ensimmäisen pahnueen syntymän ja vieroituksen väliseen porsaskuolleisuuteen (P-arvo = $2,12E-07$ ja kahdella jälkimmäisellä molemmilla $7,01E-07$), 11 (krom. 7) emakon ikään ensimmäisellä porsimisella (P-arvot $1,74E-06$ ja $3,89E-08$ välillä) ja kaksi (krom. 7) toiseen porsimisväliin (molemmilla P-arvot $4,03E-07$). Tutkimuksessa löydettiin myös suuntaa antavia (P-arvo $\leq 4,0E-06$) alueita kromosomeista 1 ja 8. Nämä kromosomialueet vaikuttavat myöhempien pahnueiden pahnuekokoon ja kuolleena syntyneiden porsaiden lukumäärään ensimmäisessä pahnueessa. Jos tässä tutkimuksessa löydetty kromosomialueet saadaan varmistettua, voidaan näitä SNP-markkereita hyödyntää merkkiavusteisen valinnan avulla kansallisessa jalostusohjelmassa. MAS:n avulla sikoja voidaan alkaa valita hedelmällisyysominaisuuksien kannalta suotuisien SNP-markkereiden suhteen ja pidemmällä tähtäimellä hedelmällisyysominaisuuksien geneettinen edistyminen alkaa nopeutua. Laajempien populaatiotutkimusten avulla voidaan myös selvittää, mitkä geenit vaikuttavat hedelmällisyysominaisuuksiin ja millaiset niiden geenivaikutukset ovat.</p>			
Avainsanat — Nyckelord — Keywords			
Yorkshire-sika, hedelmällisyysominaisuudet, SNP-markkerit, assosiaatioanalyysi			
Säilytyspaikka — Förvaringsställe — Where deposited			
Maataloustieteiden laitos, kotieläintieteen kirjasto			
Muita tietoja — Övriga uppgifter — Further information			
Tutkielman ohjaajat: erikoistutkija Pekka Uimari ja professori Jarmo Juga			

Tiedekunta/Osasto — Fakultet/Sektion — Faculty		Laitos — Institution — Department	
Faculty of Agriculture and Forestry		Department of Agricultural Sciences	
Tekijä — Författare — Author			
Sini Wallén			
Työn nimi — Arbetets titel — Title			
Whole-genome SNP association analysis of reproduction traits in the Finnish Yorkshire pig breed			
Oppiaine — Läroämne — Subject			
Animal Breeding			
Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages	
Master's thesis	May 2012	44	
Tiivistelmä — Referat — Abstract			
<p>It has been challenging to genetically improve reproduction traits with the resources of the conventional breeding because of the low heritability of these traits. In Finland genetic improvement of the reproduction traits hasn't either meet the expectations entirely. Gene markers are one possibility to make genetic improvement of the reproduction traits more effective. With an association analysis it is possible to find those gene markers which have statistically significant effect on the trait.</p> <p>The purpose of this study was to identify SNP markers associated with reproduction traits in the Finnish Yorkshire pig breed. Other purpose of this study was to find out if both Finnish pig breeds have certain chromosomal regions associated with same reproduction traits. In this study one goal was also to discover genes which are placed near statistically significant SNP markers.</p> <p>Under this study were nine female reproduction traits which are total number of piglets born in first and later parities, number of stillborn piglets in first and later parities, piglet mortality between birth and weaning in first and later parities, age at first farrowing, first farrowing interval and second farrowing interval.</p> <p>Prior to SNP association analysis, unstandardized EBVs were deregressed and corresponding weights were calculated. The association between SNP markers and deregressed EBV was studied using a mixed linear model, for each SNP separately. The model included a fixed SNP effect and a random polygenic effect. The analyses were performed using the AI-REML method. Statistical significance of the associations was based on Bonferroni-corrected P-values. In this study 20 statistically significant ($P\text{-value} \leq 2,0\text{E-}06$) associations on the reproduction traits were observed. One of these SNP markers ($P\text{-value} = 8,39\text{E-}08$, on chromosome 7) is associated in total number of piglets born in first parity, two markers ($P\text{-values} = 2,77\text{E-}07$ and $1,91\text{E-}06$, on chromosomes 7 and 1) in total number of piglets born in later parities, one marker ($P\text{-value} = 9,72\text{E-}08$, on chromosome 8) in number of stillborn piglets in first parity, three markers ($P\text{-value} = 2,12\text{E-}07$ and next two $P\text{-value}$ was $7,01\text{E-}07$, on chromosome 13) in piglet mortality between birth and weaning in first parity, 11 markers ($P\text{-values}$ between $1,74\text{E-}06$ and $3,89\text{E-}08$, on chromosome 7) in age at first farrowing and two markers ($P\text{-values} = 4,03\text{E-}07$, on chromosome 7) in second farrowing interval. In this study statistically suggestive ($P\text{-value} \leq 4,0\text{E-}06$) associations were also observed on chromosomes 1 and 8. These are associated in total number of piglets born in later parities and number of stillborn piglets in first parity.</p> <p>If these chromosomal regions found in this study are confirmed, these SNP markers will be valuable in the national breeding program through their use in marker-assisted selection. With MAS it's possible to select pigs which have favourable SNP markers as for reproduction traits. In the long run genetic improvement of the reproduction traits is going to accelerate. There is still need of more extensive population analysis in order to discover genes associated reproduction traits and to estimate those gene effects.</p>			
Avainsanat — Nyckelord — Keywords			
Yorkshire pig breed, reproduction traits, SNP markers, association analysis			
Säilytyspaikka — Förvaringsställe — Where deposited			
Department of Agricultural Sciences, library of Animal Science			
Muita tietoja — Övriga uppgifter — Further information			
Instructors: senior researcher Pekka Uimari and professor Jarmo Juga			

SISÄLLYS

Lyhenteet ja symbolit	6
1 Johdanto	7
2 Katsaus kirjallisuuteen	9
2.1 Sikojen hedelmällisyysominaisuudet	9
2.2 SNP-markkerit	12
2.3 Assosiaatioanalyysimenetelmät kotieläintutkimuksessa	15
2.4 MAS, QTL:t ja genominen valinta	16
2.5 Hedelmällisyysominaisuuksien geneettisen markkerit	17
2.6 Deregressointi	18
3 Tutkimuksen tavoitteet	19
4 Aineisto ja menetelmät	19
4.1 Suomen yorkshire-sikojen genominen aineisto	19
4.2 Sikojen hedelmällisyysominaisuudet	21
4.3 Jalostusarvojen laskenta kansallisessa jalostusarvostelussa	22
4.4 Jalostusarvojen deregressointi	22
4.5 Assosiaatioanalyysi	23
5 Tulokset	25
5.1 Aineiston tarkastelu	25
5.2 Assosiaatioanalyysi	28
5.2.1 Ensimmäisessä pahnueessa syntyneiden porsaiden lukumäärä (TNB1)	30
5.2.2 Myöhemmissä pahnueissa syntyneiden porsaiden lukumäärä (TNB2)	31
5.2.3 Ensimmäisessä pahnueessa kuolleina syntyneiden porsaiden lukumäärä (NSB1)	32
5.2.4 Myöhemmissä pahnueissa kuolleina syntyneiden porsaiden lukumäärä (NSB2)	32
5.2.5 Porsaskuolleisuus syntymästä vieroitukseen ensimmäisessä pahnueessa (PM1)	33
5.2.6 Emakon ikä ensimmäisellä porsimisella (AFF)	33
5.2.7 Toinen porsimisväli (SFI)	34
6 Tulosten tarkastelu	34

6.1 Kirjallisuudessa raportoidut QTL:t	35
6.2 Mahdollinen kandidaattigeeni	36
6.3 Vertailu maatiaisrodun tuloksiin	36
6.4. Tulosten hyödyntäminen käytännössä	37
7 Johtopäätökset	38
8 Kiitokset	38
9 Lähteet	39

LYHENTEET JA SYMBOLIT

AFF	(Age at first farrowing)	Emakon ikä ensimmäisellä porsimisella
AI-REML	(Average information restricted maximum likelihood)	Todennäköisyysfunktion maksimointiin perustuva varianssikomponenttien arvioimismenetelmä
BAC	(Bacterial artificial chromosome)	Suurikokoinen DNA-molekyyli, jonka avulla geenejä voidaan siirtää bakteereihin
CR	(Call rate)	Onnistuneesti genotyyhitettyjen SNP-merkkien osuus
EBV	(Estimated breeding value)	Jalostusarvon ennuste
FFI	(First farrowing interval)	Ensimmäinen porsimisväli
MAF	(Minor allele frequency)	Populaatiossa vähemmän esiintyvien alleelifrekvenssien suhde populaation kaikkiin alleelifrekvensseihin
MAS	(Marker assisted selection)	Merkkiavusteinen valinta
NSB1	(Number of stillborn piglets in first parity)	Kuolleina syntyneiden porsaiden lukumäärä ensimmäisessä pahnueessa
NSB2	(Number of stillborn piglets in later parities)	Kuolleina syntyneiden porsaiden lukumäärä myöhemmissä pahnueissa
PM1	(Piglet mortality between birth and weaning in first parity)	Porsaskuolleisuus syntymän ja vieroituksen välillä ensimmäisessä pahnueessa
PM2	(Piglet mortality between birth and weaning in later parities)	Porsaskuolleisuus syntymän ja vieroituksen välillä myöhemmissä pahnueissa
SFI	(Second farrowing interval)	Toinen porsimisväli
SNP	(Single nucleotide polymorphism)	Yhden emäksen monimuotoisuus
TNB1	(Total number of piglets born in first parity)	Syntyneiden porsaiden lukumäärä ensimmäisessä pahnueessa
TNB2	(Total number of piglets born in later parities)	Syntyneiden porsaiden lukumäärä myöhemmissä pahnueissa

JOHDANTO

Emakon porsastuotos on taloudellisesti ratkaiseva tekijä sikataloudelle (Rempel ym. 2010). Sianlihantuotannon taloudellisesti tärkeimpiin ominaisuuksiin kuuluu vieroitettujen porsaiden lukumäärä emakkoa kohti (Lund ym. 2002). Hedelmällisyysominaisuuksia jalostetaan Suomessa hedelmällisyysindeksin avulla. Indeksiin sisältyy syntyneiden porsaiden lukumäärä, kuolleena syntyneiden porsaiden lukumäärä, porsaskuolleisuus syntymän ja vieroituksen välillä, emakon ikä ensimmäisessä porsimisessa sekä porsimisväli (Serenius 2004). Hedelmällisyysominaisuuksien alhaisen periytymisasteen vuoksi näiden ominaisuuksien parantaminen perinteisen jalostusvalinnan avulla on osoittautunut vaikeaksi (King ym. 2003), eikä niiden perinnöllinen edistyminen ole täysin vastannut odotuksia myöskään Suomessa (FABA Palvelu 2010).

Yksi mahdollisuus hedelmällisyysominaisuuksien perinnöllisen edistymisen nopeuttamiseen on käyttää geenimerkkejä. Niitä voidaan käyttää merkkiavusteisessa valinnassa (marker assisted selection, MAS (Lande ja Thompson 1990)) tai genomisessa valinnassa (Meuwissen ym. 2001). Menetelmät poikkeavat toisistaan siten, että MAS perustuu rajattuun määrään geenimerkkejä, joiden yhteys jalostuksen kohteena olevaan ominaisuuteen on riittävällä varmuudella todettu (Lande ja Thompson 1990). Genominen valinta perustuu puolestaan koko genomin kattaviin geenimerkkeihin (Meuwissen ym. 2001), eikä genomisen valinnan yhteydessä ole tarvetta tietää yksittäisen geenimerkin yhteyttä jalostettavaan ominaisuuteen vaan valinta perustuu geenimerkkien yhteisesti antamaan ennusteeseen ja tunnettuihin geenimerkki-geeni yhteyksiin (Goddard ja Hayes 2007). Kuten edellä todettiin, MAS vaatii tiedon geenimerkin ja ominaisuuden yhteyden suuruudesta. Tämä saadaan selville assosiaatioanalyysillä. Analyysin tuloksena löydetään ne geenimerkit, joilla on tilastollisesti merkitsevä vaikutus tutkittavan ominaisuuden suhteen. Näitä geenimerkkejä voidaan käyttää MAS:ssa tai niiden avulla voidaan selvittää, mitkä geenit mahdollisesti vaikuttavat tutkittavaan ominaisuuteen. MAS ja genominen valinta ovat perinteiseen jälkeläis-, sisar- ja yksilöarvosteluun nähden erityisen tehokkaita ominaisuuksille, joiden periytymisaste on alhainen (Drogemuller ym. 2001) ja

ominaisuuksille, joiden fenotyypin mittaaminen on vaikeaa, kallista tai mahdollista vain myöhemmässä vaiheessa yksilön elämää (Lande ja Thompson 1990).

Yksi tärkeimmistä rajoitteista kotieläinten taloudellisesti tärkeiden ominaisuuksien analysoimiselle on ollut assosiaatiotutkimuksiin sopivien geenimerkkien puute (Ramos ym. 2009). Kvantitatiivisten ominaisuuksien geneettinen säätely on monimutkaista ja geneettistä muuntelua aiheuttavien geenien tunnistaminen vaatii suuren määrän geenimerkkejä esimerkiksi mikrosatelliitteja tai SNP-markkereita (Ramos ym. 2009). Viimeisten viiden vuoden aikana on tullut saataville kaupallisia koko genomin kattavia SNP-siruja, joilla voidaan samanaikaisesti testata usean kymmenen tuhannen SNP:n genotyypit. Yksi tällaisista koko genomin kattavista SNP-siruista on kansainvälisen sikagenomiikkakonsortion (International Porcine SNP chip consortium) kehittämä ja Illuminan valmistama PorcineSNP60, jolla pystytään analysoimaan 63000 SNP:tä samanaikaisesti. Uusi tekniikka on mahdollistanut entistä tehokkaamman assosiaatioanalyysin ja mahdollistanut genomisen valinnan myös sianjalostuksessa.

Tämä tutkimus on osa MTT:n Hyvinvointia ruuasta –ohjelman SikaSNiP-tutkimushanketta. SikaSNiP eli Eläinmallien ja kotieläin genomitutkimuksen hyödyntäminen elintarviketuotannossa: Sika ihmisen ravitsemuksessa –hanke on kolmivuotinen ja se keskittyy sikojen lisääntymis-, lihanlaatu-, terveys- sekä rehunkäyttöominaisuuksiin ja niiden genomiikkatutkimukseen. Hankkeen tuloksia on tarkoitus hyödyntää myöhemmin sikojen valintaohjelmassa muun muassa parantamaan sianlihan laatuominaisuuksia. Hankkeen puitteissa tehdään assosiaatioanalyysit myös sikojen hedelmällisyysominaisuuksien suhteen kummassakin kotimaisessa rodussa, yorkshire- ja maatiaisrodussa.

Tämän tutkimuksen tarkoituksena oli löytää sian hedelmällisyysominaisuuksiin vaikuttavia kromosomialueita SNP-markkereiden avulla. Löydettyjä kromosomialueita voidaan hyödyntää MAS:ssa ja niiden avulla on mahdollista paikallistaa geenejä, jotka vaikuttavat hedelmällisyysominaisuuksiin. Tutkimustulokset lisäävät parhaimmillaan tietämystä hedelmällisyysominaisuuksien geneettisestä taustasta.

2 KATSAUS KIRJALLISUUTEEN

2.1 Sikojen hedelmällisyysominaisuudet

Vieroitettujen porsaiden määrä emakkoa kohti on sianlihantuotannon taloudellisesti tärkeimpiä ominaisuuksia (Lund ym. 2002). Lundin ym. (2002) mukaan tärkeimmät tähän ominaisuuteen vaikuttavat tekijät ovat emakon ovulaatioaste, alkioden selviytyminen, sikiöiden selviytyminen porsimiseen asti ja sen aikana sekä porsaiden selviytyminen syntymästä vieroitukseen. Vieroitettujen porsaiden lukumäärää voidaan kasvattaa parantamalla mitä tahansa näistä ominaisuuksista (Lund ym 2002). Bennettin ja Leymasterin (1989) tutkimuksen mukaan pahnuekokoa voidaan kasvattaa paremmin suuntaamalla valinta ovulaatioindeksiin ja alkioden selviytymiseen tai kohdun kapasiteettiin pahnuekoon sijasta.

Pahnuekoon kasvattamisen geneettinen edistymisen on ollut hidasta (Lamberson ym. 1991), koska ominaisuus on sukupuolirajoitteinen ja sillä on lisäksi alhainen periytymisaste (Bennett ja Leymaster 1989, Lamberson ym. 1991). Tärkeimmät sikojen hedelmällisyysominaisuudet, kuten ovulaatioiden määrä, alkioden selviytyminen ja pahnuekoko, ilmenevät vain naarailla ja omaavat alhaiset periytymisasteet (Bennett ja Leymaster 1989). Hedelmällisyysominaisuuksien alhainen periytymisaste johtuu siitä, että ympäristötekijöillä kuten eläimen terveydellä ja ravitsemuksellisella tasolla on suuri vaikutus näihin ominaisuuksiin (Andersson 2001). Tämän takia näiden ominaisuuksien parantaminen jalostusvalinnan avulla on osoittautunut vaikeaksi (King 2003).

Seuraavissa kappaleissa esitetyt tarkastelut hedelmällisyysominaisuuksien jalostamisesta Suomessa perustuvat Timo Sereniuksen väitöskirjaan (Serenius 2004). Ne naaraat, jotka eivät tule kiimaan, eivät hedelmöity tai joilla on huonot emominaisuudet eli ne eivät kykene kasvattamaan jälkeläisiään, aiheuttavat porsastuotantoon ylimääräisiä kustannuksia. Kun alennetaan emakon ikää ensimmäisessä porsimisessa ja lyhennetään porsimisvälejä, saadaan porsastuotannon kustannuksia vähennetyksi. Porsasta kohti tulevia kustannuksia voitaisiin vähentää kasvattamalla vieroitettujen porsaiden lukumäärää. Vieroitettujen porsaiden lukumäärää

voidaan lisätä kasvattamalla pahnuekokoa, vähentämällä porsaskuolleisuutta ja lisäämällä emakon tuotantokauden aikana saamien pahnueiden määrää (Serenius 2004).

Tällä hetkellä Suomessa käytössä oleva hedelmällisyysindeksi sisältää syntyneiden porsaiden lukumäärän, kuolleena syntyneiden porsaiden lukumäärän, porsaskuolleisuuden syntymän ja vieroituksen välillä, emakon iän ensimmäisessä porsimisessa sekä porsimisvälin (Serenius 2004). Näistä ominaisuuksista pahnuekoon ja porsaskuolleisuuden ensimmäisen ja myöhempien pahnueiden tulokset käsitellään erillisinä ominaisuuksina. Indeksissä huomioidaan vain kaksi ensimmäistä porsimisväliä ja myös ne käsitellään erillisinä ominaisuuksina. Figen Oy:n tekemän porsastuotannon tarkkailun avulla kerätään tietoja emakon suorituskyykyyn liittyvistä ominaisuuksista. Kerättäviä tietoja ovat syntyneiden porsaiden lukumäärä, kuolleena syntyneiden porsaiden lukumäärä ja porsimispäivä (Taulukko 1.). Tietoja kerätään lisäksi vieroitettujen porsaiden lukumäärästä ja vieroituspäivistä. Myös tiedot emakoiden poistojen syistä ja poistopäivistä saadaan porsastuotannon tarkkailusta (Serenius 2004). Geneettinen edistymisen hedelmällisyysominaisuuksien suhteen ei ole ollut niin nopeaa, kuin oletettiin (FABA Palvelu 2010).

Vuosi	2006	2007	2008	2009	2010
Syntynyt porsaita/ emakko/ vuosi	24,8	25,0	25,3	25,2	26,1
Pahnuekoko syntyessä	12,1	12,2	12,3	12,4	12,5
Kuolleisuus-% syntymästä vieroitukseen	12,9	13,4	13,0	13,2	13,9
Porsimisikä, pv	369	382	384	382	380
Porsimisväli, pv	170	169	170	167	166

Taulukko 1. Vuosien 2006 – 2009 porsastuotantotulokset (Haltia 2010) sekä vuoden 2010 porsastuotantotulokset (Soili Haltia, Figen Oy, sähköpostiviesti kirjoittajalle 28.3.2012)

Aikaisempien tutkimuksien (Adamec ym. 1997 (Ref. Serenius 2004), Alfonso ym. 1997 (Ref. Serenius 2004) ja Chen ym. 2003) perusteella voidaan sanoa, että hedelmällisyysominaisuuksien periytymisasteet ovat alhaiset kaikille muille ominaisuuksille paitsi emakon ikä ensimmäisessä porsimisessa -ominaisuudelle, jonka periytymisaste on kohtalainen. Pahnuekoon periytymisaste vaihtelee 0,05 ja 0,10 välillä (Chen ym. 2003). Adamecin ym. (1997) (Ref. Serenius 2004) ja Alfonso ym. (1997) (Ref. Serenius 2004) tutkimuksien mukaan pahnuekoon periytymisaste vaihtelee 0,06 ja 0,15 välillä, porsaskuolleisuuden 0,03 ja 0,12 välillä, porsimisvälin vastaavat luvut ovat 0,04 – 0,11 ja emakon ikä ensimmäisessä porsimisessa -ominaisuuden periytymisaste on 0,32 – 0,39. Emakon ikä ensimmäisessä porsimisessa -ominaisuuden ja porsimisvälin välillä on positiivinen korrelaatio (0,25 – 0,74) (Adamec ym. (1997) (Ref. Serenius 2004). Näiden ominaisuuksien välillä vallitseva korrelaatio on hyödyllinen molemmille ominaisuuksille, mutta erityisesti juuri porsimisvälille (Serenius 2004). Syntyneiden porsaiden lukumäärän ja porsaskuolleisuuden välillä on epäsuotuisa korrelaatio. Näin ollen, jos valintaa tehdään vain pahnuekoon suhteen, porsaskuolleisuus voi lisääntyä. Vaikka hedelmällisyysominaisuuksien periytymisasteet

ovat alhaisia, tehokkaalla jalostusohjelmalla on mahdollista vaikuttaa pahnuekoon ja porsaskuolleisuuden geneettiseen edistymiseen. Valintaa täytyy tehdä vieroitettujen porsaiden suhteen joko suoraan tai epäsuorasti suuntaamalla valinta pahnuekokoon tai alhaiseen porsaskuolleisuuteen (Serenius 2004).

2.2 SNP-markkerit

SNP (single nucleotide polymorphism) on yhden emäksen monimuotoisuus. Se tarkoittaa käytännössä pistemutaatiota, jota tutkitaan uudella määrittystekniikalla. SNP-merkit kattavat koko genomin ja esiintyvät siinä hyvin tiheästi. Genotyypityksessä SNP-merkkejä on voitu hyödyntää vasta 2000-luvulla. Collins ym. (1999) mukaan SNP:tä voidaan käyttää sukupuu-, perhe- tai populaatiotutkimuksiin lajien sisällä ja etenkin silloin, kun käytetään automatisoituja matriisipohjaisia genotyypitysteknologioita. Kotieläinjalostus on ottanut SNP-merkit käyttöön viime vuosien aikana (Strauss 2010). Populaatiotutkimuksissa on tärkeää käyttää useita SNP-merkkejä, jotta saataisiin riittävä tilastollinen teho (Fries ja Durstewitz 2001).

Muihin DNA-merkkeihin verrattuna SNP:t ovat erityisen käyttökelpoisia niiden runsauden, geneettisen vakauden ja nopeiden, automatisoitujen analyysimenetelmien takia (Heaton ym. 2002). Mikrosatelliittimarkkereihin verrattuna SNP-sirujen käyttö tuottaa vähemmän kuluja genotyyppiä kohti (Garrick 2011). Kustannustehokkaat menetelmät ja SNP-merkkien kehittyminen ovat mahdollistaneet koko genomin kattavat assosiaatioanalyysit (GWAA) sekä ihmisillä että kotieläimillä (Kadarmideen ym. 2006). SNP-sirujen käyttö mahdollistaa myös useamman eläimen genotyyppittämisen kerralla, mikä lisää menetelmän tehokkuutta (Garrick 2011). Kotieläinten populaatorakenne mahdollistaa niiden käytön GWAA:ssa. Kotieläinrodut ovat muodostuneet suurempien populaatioiden jakautuessa pienempiin suljettuihin populaatioihin (Andersson 2008). Tämä on aiheuttanut rotujen sisäisen geneettisen monimuotoisuuden vähenemisen ja johtanut pitkiin haplotyyppiblokkeihin. Tämän vuoksi kotieläinten GWAA:han tarvitaan pienempi määrä SNP-merkkejä kuin esimerkiksi ihmisten vastaaviin analyysihin (Andersson 2008).

Sian genomia sekvensoidaan SGSC:n (The Swine Genome Sequencing Consortium) toimesta (Schook ym. 2005). Lähtökohtana genomien sekvensoimiselle oli hierarkkinen BAC-klooneista (bacterial artificial chromosome) koostuva shotgun-sekvenssidata, joka kattoi genomien 4-6 kertaista (Archibald ym. 2010). Lopullinen SNP-data sisälsi myös aiemmin tunnistettuja SNP:tä, jotka ovat julkisesti saatavilla eri lähteistä esimerkiksi Wageningen yliopiston tunnistamat SNP:t (Amaral ym. 2009), Illuminan 7K SNP-siru, pyrosekvensoinnin avulla tunnistetut SNP:t (Wiedmann ym. 2008), Sanger-sekvensoinnin avulla löydetty INRA:n SNP:t sekä julkisesti saatavilla olevat SNP:t, mikä sisältää myös Cambridgen yliopiston keräämät SNP:t (Ramos ym. 2009). Ensimmäinen sekvensointi valmistui 2010 heinäkuussa (Archibald ym. 2010).

Sian genomi tarkentuu koko ajan. Kun genomia sekvensoidaan, siitä julkaistaan ajoittain uudet koonnit (engl. build) (Ensembl 2009. *Sus scrofa*. Viitattu 19.10.2011). Koonti muodostuu sekvenssipaloista kootuista kromosomeista ja siitä näkyy geenien ja SNP:iden sijainti ja järjestys genomissa. Varhaisimmat koonnit sisältävät paljon virheitä, jolloin muun muassa SNP:iden paikat ja järjestys voivat olla vääriä. Uusissa koonneissa SNP:iden väliset etäisyydet saattavat muuttua, mutta SNP:iden lähiympäristö ei todennäköisesti juurikaan muutu. SNP:iden lähellä olevien geenien järjestyksessä ja sijainnissa ei pitäisi myöskään tapahtua suuria muutoksia. Uudemmissa koonneissa pyritään täydentämään mahdollisia aukkoja genomissa ja tarkentamaan genomia. Näin myös SNP:iden paikat ja järjestys tarkentuvat ja vakiintuvat. Uusi koonti on aina tarkempi ja luotettavampi kuin vanha. Tällä hetkellä käytössä on sekvenssin koonti 10.2 (Ensembl 2009. *Sus scrofa*. Viitattu 19.10.2011). Tässä tutkimuksessa käytettiin koonti 9, mutta käytön jälkeen siihenkin on tullut jo neljä päivitystä (Ensembl 2009. *Sus scrofa*, archives. Viitattu 2.2.2012). Koontiin 9 liittyy vielä monia virheitä, joten esim. parhaan SNP:n lähellä olevat geenit saattavat jonkin verran muuttua kun genomi saadaan varmemmin kuvattua.

Illuminan PorcineSNP60-sirun kehityksessä oli alun perin mukana 72 000 SNP:tä ja neljä yleisintä sikarotua: duroc, pietrain, maatiaissika, yorkshire sekä villisika (Ramos ym. 2009). Lopullinen siru sisältää 64 232 SNP-markkeria. Markkerien käyttökelpoisuus testattiin genotyypittämällä 158 yksilöä. Yksilöt edustivat samoja sikarotuja, kuin mitä sirun kehityksessäkin käytettiin. Testin perusteella selvisi, että suurin osa SNP:istä on polymorfisia (yli 57 000 SNP:tä, joiden MAF (populaatiossa

vähemmän esiintyvien alleelifrekvenssien suhde populaation kaikkiin alleelifrekvensseihin) oli suurempi kuin 0,05). SNP:iden karsimisen perusteina olivat muun muassa laatuparametrit, MAF-arvo ja sijainti genomissa. Ramosin ym. (2009) mukaan kaikissa muissa kromosomeissa paitsi X-kromosomissa SNP:iden keskimääräinen etäisyys toisistaan vaihtelee 30 ja 40 kb:n välillä. Sirun kehittämisen lisäksi tutkijat tunnistivat 372 000 uutta SNP:tä. Sirun kehittämisen aikana sian genomien sekvensointi ei ollut vielä valmistunut. Sekvensoimatta oli noin 30 % sian genomista. Puuttuvien SNP:iden sijainnit genomissa arvioitiin ihmis-sika –kartan avulla vertailemalla sekä hyödyntämällä BAC-kontigeista koostuvaa karttaa (Ramos ym. 2009).

Assosiaatioanalyysissä tärkeimpiä asioita käytettävän sirun kannalta ovat SNP:iden lukumäärä ja niiden keskimääräinen etäisyys toisistaan sekä tutkittavassa populaatiossa vallitseva kytkentäepätasapaino. Anderssonin ym. (2008) mukaan populaatiossa vallitsevan kytkentäepätasapainon (LD) tyyppi määrää assosiaatioanalyysiin tarvittavien SNP-merkkien lukumäärän (Andersson 2008). Eri tutkimuksien mukaan yleisimmillä sikaroduilla, kuten durocilla, maatiaisella ja yorkshirella, LD:tä esiintyy 40 – 60 kb:n (Jungerius ym. 2005) ja jopa 400 kb:n (Amaral ym. 2008) pituisilla alueilla. Tutkijat ovat sitä mieltä, että sian koko genomien kattavia assosiaatioanalyysyjä varten markkeritiheyden tulisi olla 5 – 10 markkeria/ cM (Amaral ym. 2008 ja Du ym. 2007).

Koko genomien kattava SNP-genotyypitys tuli mahdolliseksi sialla vuonna 2009 (Ramos ym. 2009). Uusi SNP-teknologia antaa hyödyllisiä työkaluja, joilla voidaan tutkia sikapopulaatioiden geneettistä rakennetta ja se myös sallii yksityiskohtaisemman vertailun populaatioiden välillä kuin aiemmat sukupuutietoihin perustuvat lähestymistavat (Uimari ja Tapio 2011). Keskimääräinen kytkentäepätasapaino (LD) PorcineSNP60-sirun vierekkäisten SNP:iden välillä oli 0,46 suomalaisella yorkshire-rodulla ja 0,43 maatiaisrodulla (Uimari ja Tapio 2011). Uimarin ja Tapion (2011) tutkimuksen mukaan suomalaisen yorkshire-rodun tehollinen populaatiokokoo on 61 ja vastaava luku maatiaisrodulle on 91. Suomalaisessa yorkshire-populaatiossa vallitsi vahvempi kytkentäepätasapaino kuin maatiaissialla (Uimari ja Tapio 2011). LD-arvo suomalaiselle yorkshirelle oli $r^2 = 0,16$ (kun SNP:t olivat 3 Mb:n etäisyydellä toisistaan) ja vastaava arvo maatiaisrodulle oli 0,12. Koska molempien rotujen LD-arvot olivat kohtuullisia, koko genomien kattaviin SNP-merkkeihin perustuvan valinnan eli

genomisen valinnan odotetaan olevan tehokas kummallekin rodulle (Uimari ja Tapio 2011).

2.3 Assosiaatioanalyysimenetelmät kotieläintutkimuksessa

Populaation assosiaatiotutkimuksien tavoitteena on tunnistaa polymorfiat, joiden esiintyvyys vaihtelee esimerkiksi erilaisissa sairaustilanteissa olevien yksilöiden välillä (Balding 2006). Assosiaatiotutkimuksien avulla voidaan tällaisessa tilanteessa saada selville tautiriskiä lisäävien alleelien tai suojaavien alleelien vaikutukset (Balding 2006). Tällä hetkellä markkinoilla olevien 50 000 – 60 000 SNP-sirujen tarkkuus riittää assosiaatiotutkimuksiin, jos vähintään 60 % markkereista on informatiivisia ja hyvälaatuisia populaatiotutkimuksissa (Schulman ym. 2011).

Assosiaatio-menetelmä hyödyntää koko populaation LD:n perheiden sisäisen LD:n sijasta (Schulman ym. 2011). Assosiaatioanalyysimenetelmiä ovat esimerkiksi kandidaatti-polymorfia, kandidaattigeeni, hienokartoitus sekä koko genomien kattava assosiaatiomenetelmä (Balding 2006). Kandidaatti-polymorfia keskittyy yksittäisiin polymorfioihin, kandidaattigeeni-menetelmässä tarvitaan muutamia kymmeniä SNP:tä geeniä kohti, hienokartoitukseen riittää muutama sata SNP, kun taas viimeiseen menetelmään tarvitaan kymmeniä tuhansia SNP:tä.

Yksinkertainen tapa testata populaatiossa vallitseva assosiaatio on käyttää yksittäistä markkeri-regressiota, jossa fenotyyppi on regressoitu kiinteällä SNP-merkin genotyypillä (Schulman ym. 2011). Jos yksilöt tulevat eri alapopulaatioista, ja/ tai ovat läheisessä sukulaisuudessa toisiinsa, tarvitaan edistyneempiä menetelmiä, jotka ottavat huomioon populaation rakenteen (Voight ja Pritchard 2005). Näin voidaan välttää väärät positiiviset assosiaatiotulokset (Voight ja Pritchard 2005). Yksi tapa poistaa populaation rakenteen vaikutukset ja perheiden sukulaisuus on käyttää sekamallia (Yu ym. 2006). Assosiaatiotutkimuksissa, jotka perustuvat LD:hen, tarvittava tilastollinen teho ilmenee vain QTL:n läheisyydessä, joten markkerien tulisi sijaita riittävän tiheästi genomissa (Simianer ja Pimentel 2009). SNP:t ovat yleisin geneettisen muuntelun aiheuttaja ja ne ovat erityisen hyödyllisiä assosiaatiotutkimuksiin ja kytöntäkarttoituksiin sialla (Li ym. 2008). Kotieläinten genotyypittäminen SNP-

markkereiden suhteen on luonut tarpeen uusien tilastollisten menetelmien kehittämiseksi, erityisesti QTL-kartoituksen tarpeisiin (Simianer ja Pimentel 2009).

2.4 MAS, QTL:t ja genominen valinta

Merkkiavusteisella valinnalla (MAS) voidaan tehdä valintaa, molemmilla sukupuolilla, pahnuekokoon liittyvien ominaisuuksien suhteen jo hyvin varhaisessa vaiheessa yksilön elämää ja näin tehostaa valintaa (Cassadyn ym. 2001). Ominaisuudet, joihin vaikuttaa pieni määrä suurivaikutteisia geenejä, ovat soveltuvampia MAS:hin kuin ne ominaisuudet, joihin vaikuttaa suuri määrä pienivaikutteisia geenejä (Wilkie ym. 1999). Landen ja Thompsonin (1990) mukaan MAS:n tehokkuus fenotyyppisessä valinnassa on suurimmillaan sukupuolirajoitteisissa ominaisuuksissa tai ominaisuuksissa, joilla on alhaiset periytymisasteet. MAS:n käyttö hedelmällisyysominaisuuksille, kuten pahnuekoolle, vaikuttaisi erityisen lupaavalta, mikä johtuu näiden ominaisuuksien alhaisista periytymisasteista ja sopivien geneettisten markkerien olemassa olostä (Drogemuller ym. 2001) sekä sukupuolirajoittuneisuudesta ja ominaisuuden ilmenemisestä myöhään yksilön elämässä (Lande ja Thompson 1990).

Geneettisten markkerien informaatiota voitaisiin käyttää hyväksi niiden urosten valitsemisessa, jotka kantavat naarashedelmällisyyden kannalta toivottuja alleleja (Drogemuller ym. 2001 ja King ym. 2003). Näin voitaisiin valita sopivat eläimet jo ennen kuin ne saavuttavat sukukypsyyden ja saavat ensimmäisen pahnueensa (Drogemuller ym. 2001 ja King ym. 2003). Tämä johtaa tehokkuuden parantumiseen, sukupolvien välin lyhentymiseen ja ominaisuuden geneettisen edistymisen nopeutumiseen (Drogemuller ym. 2001). Perinteisiä menetelmiä käytettäessä hedelmällisyysominaisuuksien valinta on paljon tehottomampaa kuin kasvu- tai teurasominaisuuksien valinta (Wilkie ym. 1999).

Garrickin (2011) mukaan sekä perinteisissä QTL-tutkimuksissa että genomisessa valinnassa vaikutusten havaitseminen riippuu segregoituvan markkerigenotyypin ja kausaalisen polymorfian välisestä yhteydestä, jonka voimakkuus heijastaa kytkeäntäpätasapainon (LD) voimakkuutta. Caluksen ym. (2008) mukaan genominen valinta perustuu oletukseen, että geneettisten merkkien välillä vallitseva LD on

samansuuruinen sekä referenssipopulaatiossa että tutkittavassa populaatiossa. Genominen valinta ei vaadi sukupuutietoja, mutta hyöttyy SNP-markkereiden tiheästä esiintyvyydestä ja siitä, että fyysisesti lähellä toisiaan sijaitsevilla lokuksilla on voimakkaampi LD kuin kaukana toisistaan sijaitsevilla lokuksilla (Garrick 2011). Jos oletetaan, että genomi on kyllästetty SNP-markkereilla, jokaisen QTL:n pitäisi olla lähellä jotain genotyyppitettyä SNP:tä ja toivottavasti vähintään yksi SNP olisi riittävässä kytkentäepätasapainossa jonkin QTL:n kanssa (Garrick 2011). Perinteiseen valintaan verrattuna genominen valinta on virheettömämpää, erityisesti ominaisuuksille, joilla on alhaiset periytymisasteet (Calus ym. 2008).

Genominen valinta soveltuu hyvin hedelmällisyysominaisuuksille (Meuwissen ym. 2001), sillä genomisessa valinnassa käytetään hyödyksi tiheästi sijaitsevien markkerien antamaa genotyyppi-infomaatiota, jotta valintakandidaateille voidaan laskea EBV:t (Schulman ym. 2011). Strandénin ja Garrickin (2009) mukaan genominen valinta sisältää arvion SNP-genotyyppienä tunnettujen genomialueiden geneettisestä arvosta. Tätä tietoa voidaan käyttää hyväksi ennustettaessa yksittäisten eläinten geneettinen arvo niiden perimien genomialueiden perusteella (Strandén ja Garrick 2009). Schulmanin ym. (2011) mukaan genomisessa valinnassa ei edellytetä tietoa geeneistä, jotka vaikuttavat tutkittavaan ominaisuuteen. Hedelmällisyysominaisuuksiin assosioituneiden geenien löytäminen voi kuitenkin antaa arvokasta tietoa monimutkaisesta geneettisestä verkosta, joka aiheuttaa vaihtelua eläinten lisääntymisbiologiaan (Schulman ym 2011).

2.5 Hedelmällisyysominaisuuksien geneettisen markkerit

Eri sikaroduissa nähtävät erot hedelmällisyysominaisuuksien ilmenemisessä osoittavat hyödyllisen geneettisen muuntelun olemassaolon (King ym. 2003). Sikojen naarashedelmällisyys-ominaisuuksien geneettisen säätelyn ymmärtäminen voi tarjota mahdollisuuden luonnollisen vaihtelun hyödyntämiseen ja valikoivien jalostusohjelmien parantamiseen MAS:n avulla (King ym. 2003). Erilaisia geneettisiä markkereita on tullut vuosien kuluessa saataville: kausatiiviset mutaatiot (suoranaiset markkerit), linkitetyt markkerit eli markkerin ja QTL:n välillä vallitsee kytkentäepätasapaino (LD-markkerit) sekä sellaiset linkitetyt markkerit, että markkerin ja QTL:n välillä vallitsee kytkentätasapaino (LE-markkerit) (Dekkers 2004). LD-markkerit voidaan tunnistaa

käyttämällä kandidaattigeeni-menetelmää (Dekkers 2004 (Ref. Rothschild ja Soller 1997)) tai hienokartoitusta (Andersson 2001).

Lisääntymisominaisuuksien geneettisten markkerien tunnistamiseen on käytetty kahta menetelmää (King ym. 2003). Ensin tehdään genomiskannaus DNA-markkereiden avulla, jotta voitaisiin tunnistaa ominaisuuksiin vaikuttavat QTL:t (Rathje ym. 1997, Rohrer ym. 1999, Wilkie ym. 1999). Seuraavaksi käytetään fysiologista kandidaattigeeni-lähestymistapaa, joka hyödyntää polymorfioita (Drogemuller ym. 2001, Jiang ym. 2001, Rothschild ym. 1996). Hyödylliset polymorfiat esiintyvät geeneissä tai lähellä geenejä ja niiden tiedetään, hedelmällisyyden assosiaatiotestien perusteella, vaikuttavan lisääntymisen aikana (Drogemuller ym. 2001, Jiang ym. 2001, Rothschild ym. 1996).

2.6 Deregressointi

Ostersenin ym. (2011) tutkimuksessa saatiin genomisille jalostusarvoille 18 - 39 % korkeammat luotettavuudet käytettäessä deregressoituja EBV:itä verrattuna EBV:ihin (jalostusarvon ennusteisiin). Puhdasrotuisten sikojen datassa EBV:illä on yleensä alhaiset ja vaihtelevat luotettavuudet (Ostersen ym. 2011). Yleensä vain osa sioista on genotyyppitetty ja nämä eläimet ovat usein läheistä sukua toisilleen (Ostersen ym. 2011). Tämän takia tarvitaan vastaava muuttuja ja tilastollinen menetelmä, jotka ovat sopivia tällaisen datan käsittelemiseen (Ostersen ym. 2011). Garrickin ym. (2009) mukaan tähän päivään mennessä EBV ja deregressoitu EBV ovat kaikkein lupaavimpia mahdollisten vastaavien muuttujien joukosta.

Deregressoitu EBV, josta on poistettu sukulaisista johtuva informaatio ja sen määrä, antaa tarkemmat genomiset jalostusarvot kuin EBV (Garrick ym. 2009). Tähän on kaksi syytä. Ensiksi, deregressoidun EBV:n käyttö vastaavana muuttujana aiheuttaa vähemmän kahteen kertaan laskemista verrattuna EBV:hen, koska deregressoitu EBV sulkee pois perityn informaation (Garrick ym. 2009). Tällöin deregressoitu EBV sisältää vain eläimen oman ja sen jälkeläisen informaation (Garrick ym. 2009). Toiseksi, kun käytetään EBV:tä vastaavana muuttujana, aiheuttaa se genomisten jalostusarvojen kaksinkertaisen pienentymän, varsinkin silloin kun EBV:n luotettavuudet ovat alhaiset

(Garrick ym. 2009). Deregressointi eliminoi EBV:n sisältämän kutistumisen ja siksi deregressoitu EBV käyttäytyy kuin havaintojen periytymisasteet olisivat yhtä suuria kuin deregressoitujen EBV:iden luotettavuudet (Garrick ym. 2009).

3 TUTKIMUKSEN TAVOITTEET

Tutkimuksen tavoitteena oli löytää ne kromosomialueet, jotka ovat yhteydessä yorkshire-sian hedelmällisyysominaisuuksiin. Tarkoituksena oli myös selvittää, ovatko molemmilla suomalaisilla roduilla tietyt kromosomialueet yhteydessä samoihin hedelmällisyysominaisuuksiin. Lisäksi tutkimuksessa pyrittiin löytämään tilastollisesti merkitsevien SNP-markkereiden lähellä sijaitsevia geenejä.

4 AINEISTO JA MENETELMÄT

4.1 Suomen yorkshire-sikojen genominen aineisto

Tutkimuksen fenotyyppiaineisto hankittiin suomalaisilta porsastuotannon tarkkailussa mukana olleilta tiloilta. Tarvittavat tiedot saatiin suoraan tilojen tarkkailutiedoista. Tutkimuksen aineistoon otettiin vain puhdasrotuisia yorkshire-sikoja. Tutkimuksen alkuperäinen aineisto koostui 316 suomalaisen yorkshire-rotuisen keinosiemennyskarjun tiedoista. Kaikille karjuille oli laskettu kansalliset jalostusarvon ennusteet (EBV:t). Assosiaatioanalyysiin valittiin vain ne karjut, joiden call rate (CR) oli 90 tai sen yli. Call rate-luku kertoo SNP-merkkien genotyypityksen onnistumisesta. Se osuus eläimen SNP-merkeistä, joka on saanut AA, AB tai BB-genotyypin kertoo eläimen CR:n (Kaava 1).

$$CR = \frac{(N - n_{miss})}{N} \quad (\text{Kaava 1})$$

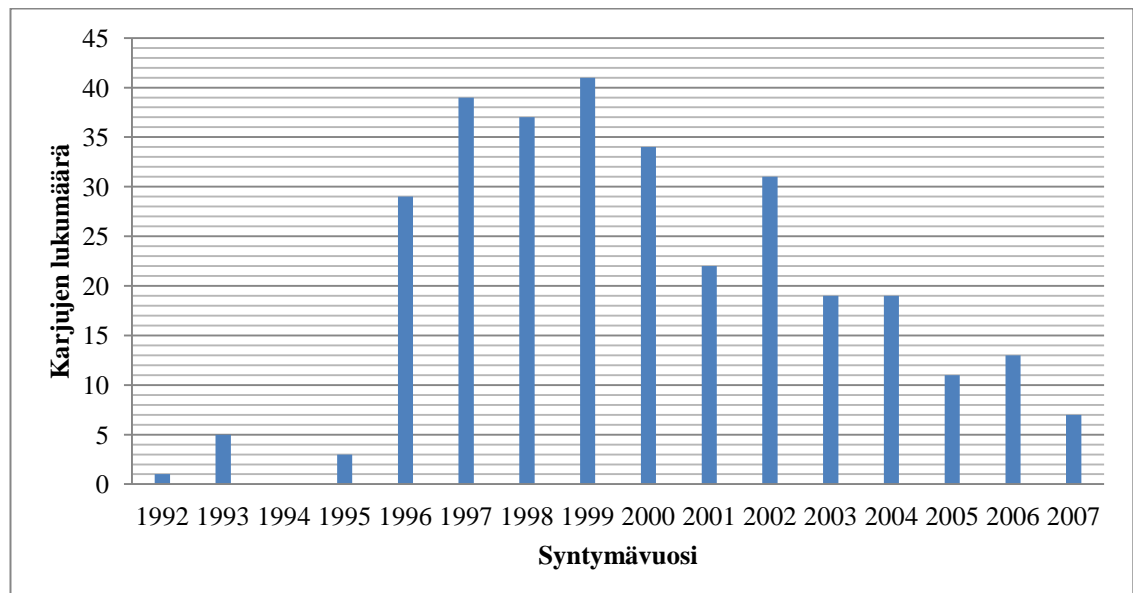
jossa N on kaikkien SNP-merkkien lukumäärä ja n_{miss} on epäonnistuneiden SNP-merkkien lukumäärä eli ne SNP:t, jotka eivät saaneet analyysissä mitään genotyyppiä. Seitsemän karjua jouduttiin poistamaan assosiaatioanalyysistä, koska niiden CR:t olivat alle 90. SNP:ille laskettiin myös HW- ja MAF-arvot (Kaavat 2 ja 3), mutta niiden perusteella ei tehty karsintaa. HW-arvo eli χ^2 -arvo lasketaan havaittujen ja

odotettujen genotyypifrekvenssien avulla. HW-frekvenssit tulevat alleelifrekvensseistä satunnaisten paritumisen perusteella. Jos saatu HW-arvo on pienempi kuin χ^2 -taulukosta saatu arvo, SNP:n genotyypit ovat jakautuneet Hardy-Weinbergin tasapainolain mukaisesti. MAF-arvo (minor allele frequency) saadaan populaatiossa vähemmän esiintyvien alleelifrekvenssien ja populaation kaikkien alleelien frekvenssien suhteena (Kaava 3). Assosiaatioanalyysin tuloksena saatujen tilastollisesti merkitsevien SNP-merkkien CR-, HW- ja MAF-arvot tarkastettiin.

$$HW \text{ eli } \chi^2 - \text{arvo} = \frac{(\text{hav. frek.} - \text{odot. frek.})^2}{\text{odot. frek.}} \quad (\text{Kaava 2})$$

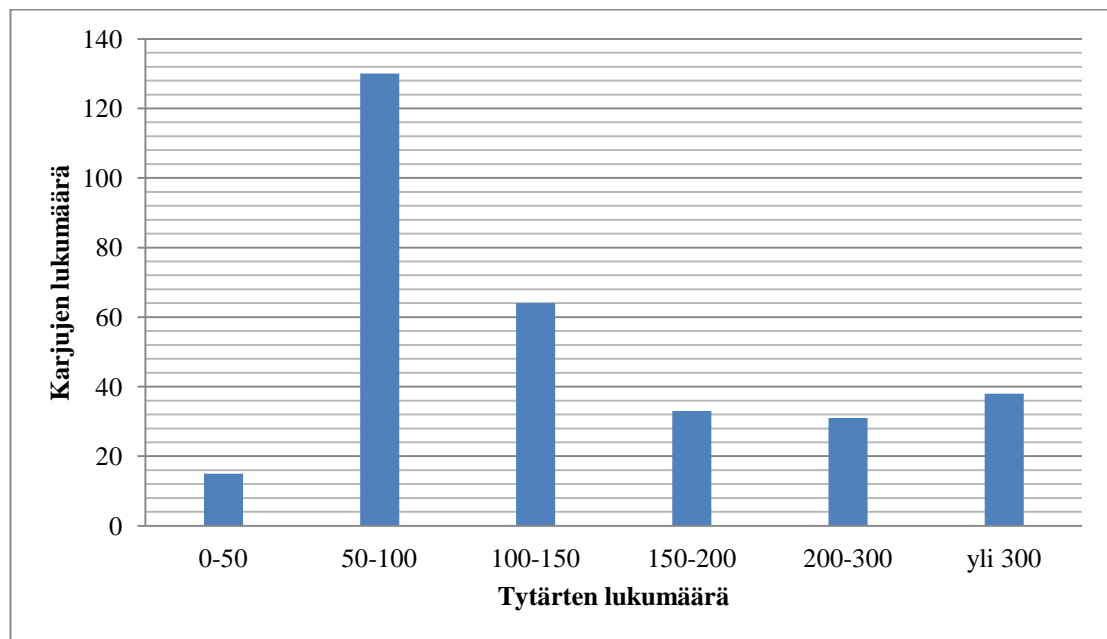
$$MAF = \frac{\text{ma. frek.}}{\text{alleelien kok. frek.}} \quad (\text{Kaava 3})$$

Kaksi karjua oli genotyyppitetty kahteen kertaan, joten eläimet esiintyivät myös tiedostossa kahteen kertaan. Nämä eläimet poistettiin analyysistä. Lopullinen assosiaatioanalyysiin sopiva aineisto koostui 307 karjun tiedoista. Karjut olivat syntyneet vuosien 1992 ja 2007 välillä (Kuva 1).



Kuva 1. Karjujen syntymävuodet.

Karjujen keskimääräinen tytärten lukumäärä oli 159. 15 nuoremmalla (2002 - 2007 syntyneellä) karjulla oli alle 50 tytärtä ja 38 vanhemmalla (suurin osa syntynyt 1992 - 2000) karjulla oli yli 300 tytärtä (Kuva 2). Suurin tytärten lukumäärä karjua kohti oli 1 473.



Kuva 2. Karjujen tytärten lukumäärä.

DNA:t eristettiin 304 karjulta karvanäytteestä ja 7 karjulta spermasta. Eristämiseen käytettiin Qiagenin Dneasy Blood & Tissue-kittiä. DNA:n tavoitekonsentraatiot olivat 100 ng/ µl spermanäytteistä ja 50 ng/ µl karvanäytteistä. Genotyypitykseen käytettiin 20 µl DNA:ta jokaisesta näytteestä. Genotyypitys tehtiin Suomen molekyyli lääketieteen instituutin (FIMM) toimesta käyttäen PorcineSNP60-sirua.

4.2 Sikojen hedelmällisyysominaisuudet

Tutkimuksessa tarkasteltiin yhdeksää naarashedelmällisyyden ominaisuutta. Ominaisuudet olivat syntyneiden porsaiden lukumäärä ensimmäisessä pahnueessa eli TNB1 (total number of piglets born in first parity), syntyneiden porsaiden lukumäärä myöhemmissä pahnueissa eli TNB2 (total number of piglets born in later parities), kuolleena syntyneiden porsaiden lukumäärä ensimmäisessä pahnueessa eli NSB1

(number of stillborn piglets in first parity), kuolleena syntyneiden porsaiden lukumäärä myöhemmissä pahnueissa eli NSB2 (number of stillborn piglets in later parities), porsaskuolleisuus syntymästä vieroitukseen ensimmäisessä pahnueessa eli PM1 (piglet mortality between birth and weaning in first parity), porsaskuolleisuus syntymästä vieroitukseen myöhemmissä pahnueissa eli PM2 (piglet mortality between birth and weaning in later parities), emakon ikä ensimmäisellä porsimisella eli AFF (age at first farrowing), ensimmäinen porsimisväli eli FFI (first farrowing interval) sekä toinen porsimisväli eli SFI (second farrowing interval).

4.3. Jalostusarvojen laskenta kansallisessa jalostusarvostelussa

Jokaiselle yhdeksälle hedelmällisyysominaisuudelle oli laskettu EBV. EBV:iden laskentaan käytettiin kansallisessa arvostelussa kahta eri mallia. TNB1-, TNB2-, NSB1-, NSB2-, PM1- ja PM2-ominaisuuksille käytetyssä lineaarisessa mallissa kiinteinä tekijöinä olivat karja-vuosi-, vuosi-kuukausi-, hedelmöitystyyppi-, pahnueen rotu- sekä emakon ikä porsimassa –tekijät. Mallin satunnaistekijät koostuivat pahnueen karjun vaikutuksesta, pysyvistä ympäristötekijöistä sekä eläintekijöistä. Ominaisuuksissa TNB2, NSB2 ja PM2 huomioitiin myös se, monesko pahnue oli kyseessä. AFF-, FFI- ja SFI-ominaisuuksien lineaarisessa mallissa kiinteinä tekijöinä olivat karja-vuosi -ja karja-kuukausi -tekijät sekä satunnaistekijänä eläintekijä. Ominaisuudet FFI ja SFI sisälsivät myös emän rodun vaikutuksen.

4.4 Jalostusarvojen deregressointi

Schaefferin (2006) mukaan jalostusarvojen ennustaminen genomisen tiedon pohjalta tapahtuu kaksivaiheisesti. Ensin pyritään arvioimaan pienten genomialueiden tai merkkien vaikutus referenssipopulaatiosta (Schaeffer 2006). Tässä käytetään apuna populaation yksilöiden markkerigenotyypppejä ja regressioanalyysiä (Garrick ym. 2009). Tässä vaiheessa voidaan käyttää hyödyksi myös fenotyyppi-informaatiota, toistuvia havaintoja, läheisten perheenjäsenten tietoja esimerkiksi jälkeläisinformaatiota, jalostusarvon ennusteita (EBV:t) tai näiden deregressoituja vastineita (Garrick ym. 2009). Seuraavaksi valintakandidaattien genominen arvo ennustetaan niiden genotyyppien ja esianalyysissä saatujen tietojen perusteella (Schaeffer 2006). Jos

esianalyysissä käytetty data koostuu EBV:istä, yksilöillä on yleensä vaihtelevat jalostusarvojen r^2 -arvot eli luotettavuudet. Tämä ongelma voidaan välttää deregressoimalla EBV:t eli jakamalla EBV:t niiden r^2 -arvoilla (Garrick ym. 2009).

Tässä tutkimuksessa käytettiin Garrickin ym. (2009) menetelmän mukaista EBV:iden deregressointia ja painokertoimien laskemista. Painokertoimet määrytyivät jalostusarvojen takana olevan informaation perusteella. Laskemisen perusteena käytettiin yksilön omia ja sen vanhempien EBV:itä ja niiden r^2 -arvoja. Garrickin ym. (2009) mukaan kaikki geneettinen varianssi ei selity markkereilla. Tähän on kaksi syytä: markkerit eivät peitä koko genomia eikä markkereiden ja niihin yhteydessä olevien geenien välinen kytkentäepätasapaino ole aina täydellistä (Garrick ym. 2009). Tutkimuksessa käytettiin Garrickin ym. (2009) menetelmän mukaista parametria c , mikä kuvaa sen varianssin osuutta, joka ei selity markkereilla. Käytetty menetelmä olettaa, että yksilön vanhemmat eivät ole sukua toisilleen eivätkä sukusiitettyjä (Garrick ym. 2009). Tässä tutkimuksessa suurin osa keinosiemennyskarjujen EBV:iden informaatiosta tulee jälkeläisiltä eikä vanhemmilta, joten ko. oletus vaikuttaa hyvin vähän tutkimuksen assosiaatioanalyysiin.

Kun EBV:t deregressoidaan, poistetaan samalla myös sukulaisista johtuva informaatio (Garrick ym. 2009). Sukulaisista saatavan informaation poistamiseen on useita syitä. Ensiksi, eläimille voidaan laskea EBV, vaikka näillä ei olisi omia eikä jälkeläistensä havaintoja (Garrick ym. 2009). Tällaisista eläimistä voidaan hyödyntää vain niiden vanhempien ja eläinten omien EBV:iden antama informaatio. Lisäksi, jos vanhemmat periivät suurivaikutteisen geenin, puolet jälkeläisistä saa suotuisan alleelin ja puolet epäsuotuisan alleelin. Kuitenkin molempien jälkeläisryhmien EBV:t supistetaan kohti vanhempien keskiarvoa (Garrick ym. 2009).

4.5 Assosiaatioanalyysi

SNP-markkereiden ja deregressoitujen EBV:iden välisen assosiaation selvittämiseen käytettiin lineaarista sekamallia, joka sisälsi kiinteän SNP-vaikutuksen ja satunnaisten polygeenisen vaikutuksen.

Käytetty malli: $y_i = \mu + b * x_i + a_i + e_i$,

jossa y_i on deregressoitu EBV; μ on yleiskeskisarvo; b on regressiokerroin; x_i kuvaa sitä, onko yksilöllä 0 1 vai 2 kopiota alleelistä, jota populaatiossa esiintyy vähemmän; a_i on satunnainen polygeeninen vaikutus oletuksella $a_i \sim N(0, A\sigma_a^2)$, missä A on sukulaisuusmatriisi ja σ_a^2 polygeeninen varianssi; e_i on satunnainen residuaalivaikutus oletuksella $e_i \sim N(0, I\sigma_e^2/w)$, missä I on identiteettimatriisi, σ_e^2 residuaalivarianssi ja w painokerroin, joka laskettiin Garrickin ym. (2009) menetelmän mukaisesti. Assosiaatioanalyysi tehtiin AI-REML-menetelmällä DMU-ohjelmalla (Madsen ym. 2006). REML-menetelmää käytetään varianssikomponenttien analysoimiseen eläinjalostussovelluksissa (Patterson ja Thompson 1971). Ominaisuuksien tilastollisesti merkitsevien SNP-markkereiden läheisyydessä sijaitsevien QTL:ien etsintään käytettiin sian QTL-tietokantaa (NAGP Bioinformatics team 2003. PigQTLdb. Viitattu 28.3.2012) ja ominaisuuksien lähellä olevien geenien tarkasteluun sian sekvenssin koontia 9 (Ensembl 2009. Sus scrofa. Viitattu 19.10.2011) ja NCBI-tietokantaa (National Center for Biotechnology Information 1993. NSBI. Viitattu 17.8.2012). Aineiston tilastollinen tarkastelu ja perustunnuslukujen laskenta tehtiin SAS 9.2 –ohjelmalla (SAS Institute Inc, Cary, NC).

Tutkimuksen analyysin tilastollinen merkitsevyys perustui Bonferroni-korjattuihin P-arvoihin. Bonferronin korjaus käsittelee yksittäisiä testejä toisistaan riippumattomina. Tarkoituksena oli, että väärin positiivisten tulosten tasona olisi 0,05. Lisäksi oletettiin, että 50 000 tuloksesta 25 000 tulosta olisi riippumatonta.

$$P - arvo: \frac{0,05}{50\,000} = 1,0E - 06 \quad (\text{Kaava 4})$$

$$P - arvo: \frac{0,05}{25\,000} = 2,0E - 06 \quad (\text{Kaava 5})$$

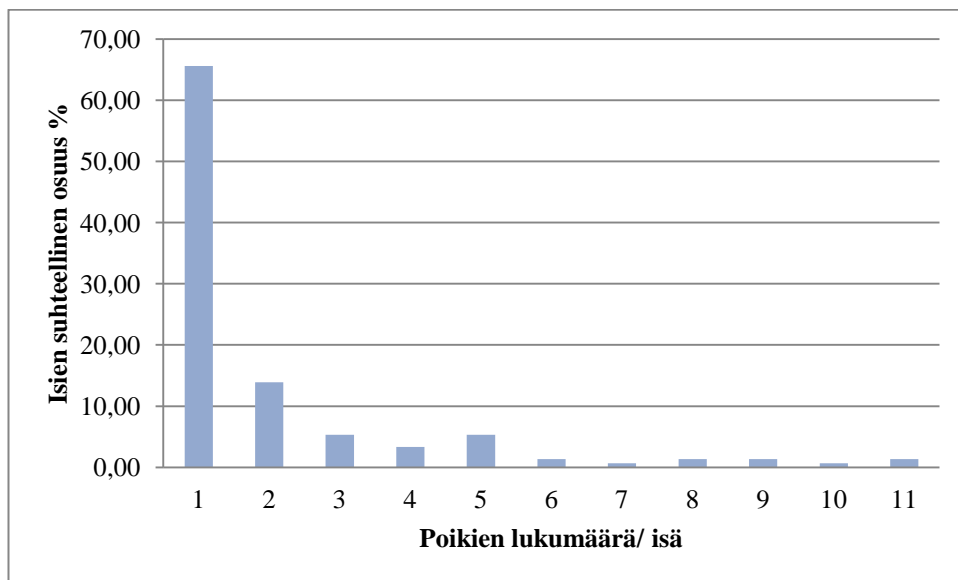
$$P - arvo: \frac{0,1}{25\,000} = 4,0E - 06 \quad (\text{Kaava 6})$$

Tällöin P-arvon pitäisi olla 1,0E-06 ja 2,0E-06 välillä (Kaavat 4 ja 5). Tässä tutkimuksessa tilastollisesti merkitsevän tuloksen P-arvoksi määriteltiin 2,0E-06 tai vähemmän ja suuntaa-antavan tuloksen P-arvoksi 4,0E-06 tai vähemmän (Kaava 6).

5 TULOKSET

5.1. Aineiston tarkastelu

Aineisto koostui 310 yorkshire-rotuisesta karjusta, joilla oli yhteensä 151 eri isää. Aineistossa kahdella isällä oli 11 poikaa ja 98 isällä oli vain yksi poika (Kuva 1). Koska rotu on pidetty suurimmilta osin suljettuna, on oletettavaa, että aineistossa olevat karjut ovat kaikki sukua toisilleen.



Kuva 1. Karjujen isäjakauma.

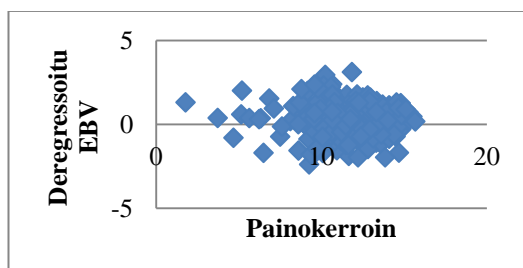
Alkuperäisen aineiston kolmen karjun call rate -arvot olivat alle 0,90 ja ne poistettiin analyysistä. Loppujen genotyyppitettyjen karjujen keskimääräinen call rate oli 0,95. Suurimmalla osalla karjuista call rate -arvot vaihtelivat 0,90 ja 0,95 välillä (Taulukko 1).

Taulukko 1. Karjujen call rate -arvojen jakauma.

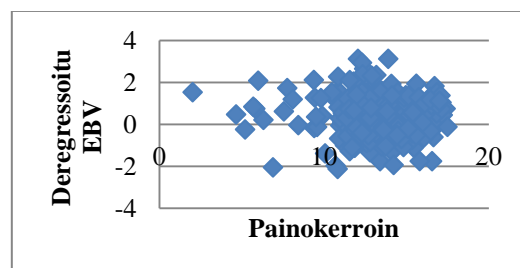
Absoluuttinen frekvenssi = karjujen lukumäärä

Call rate	Absoluuttinen frekvenssi	Suhteellinen frekvenssi	Summa-frekvenssi
0.95 - 1.0	309	0.978	0.978
0.90 - 0.95	4	0.013	0.991
0.85 - 0.90	2	0.006	0.997
0.80 - 0.85	1	0.003	1.000
0.00-0.80	0	0.000	1.000
Yhteensä	316		

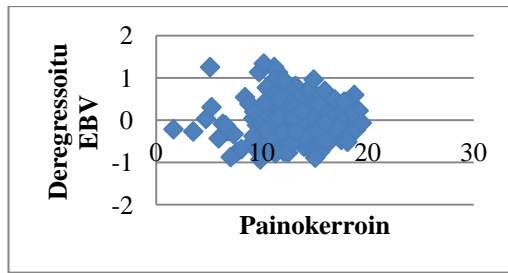
Seuraavissa kuvissa (Kuvat 2 - 10) on jokaisen ominaisuuden deregressoidut EBV:t ja niitä vastaavat painokertoimet. Ensimmäisen (FFI) ja toisen porsimisvälin (SFI) (Kuvat 9 ja 10) EBV:iden painokertoimien alarajaksi on asetettu nolla. Kuvassa 2 on ensimmäisessä pahnueessa syntyneiden porsaiden lukumäärä (TNB1), kuva 3: myöhemmissä pahnueissa syntyneiden porsaiden lukumäärä (TNB2), kuva 4: ensimmäisessä pahnueessa kuolleena syntyneiden porsaiden lukumäärä (NSB1), kuva 5: myöhemmissä pahnueissa kuolleena syntyneiden porsaiden lukumäärä (NSB2), kuva 6: ensimmäisen pahnueen syntymän ja vieroituksen välinen porsaskuolleisuus (PM1), kuva 7: myöhempien pahnueiden syntymän ja vieroituksen välinen porsaskuolleisuus (PM2), kuva 8: emakon ikä ensimmäisellä porsimisella -ominaisuus (AFF), kuva 9: ensimmäinen porsimisväli (FFI) ja kuva 10: toinen porsimisväli (SFI).



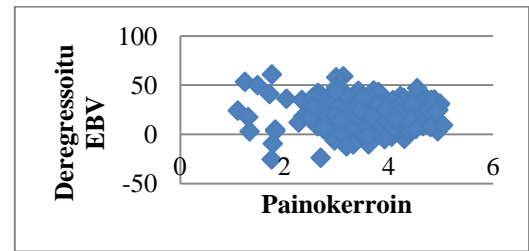
Kuva 2. Deregressoitujen EBV:iden ja painokertoimien välinen jakauma (TNB1).



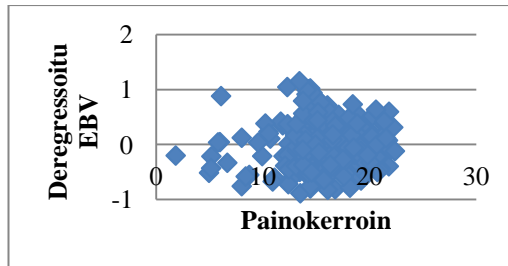
Kuva 3. TNB2



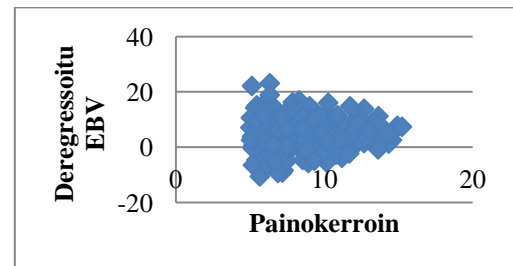
Kuva 4. NSB1



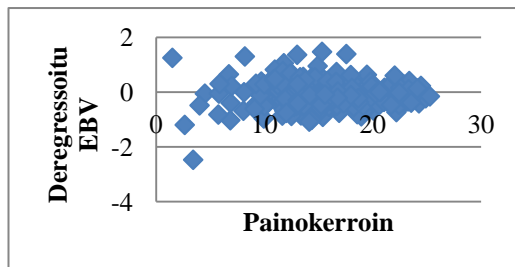
Kuva 8. AFF



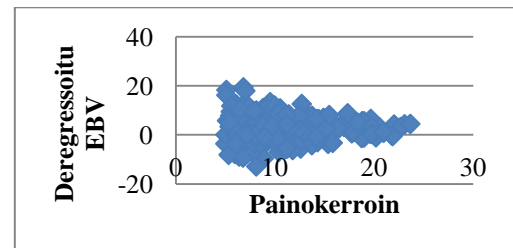
Kuva 5. NSB2



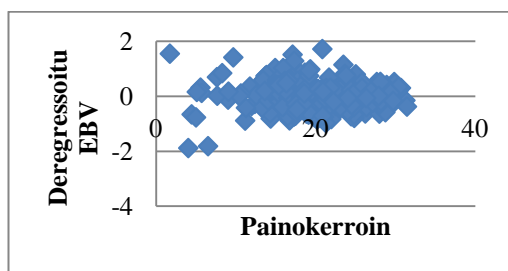
Kuva 9. FFI



Kuva 6. PM1



Kuva 10. SFI



Kuva 7. PM2

Eri ominaisuuksien karjujen lukumäärät vaihtelivat 268 ja 309 välillä. Vähiten karjuja eli havaintoja oli toiselle porsimisvälille (Taulukko 2).

Taulukko 2. Deregressoitujen EBV:iden tilastolliset tunnusluvut.

Ominaisuus, yksikkö	N	μ	σ	Min. - max.
TNB1, porsas	309	0,28	0,90	-2,40 – 3,11
TNB2, porsas	309	0,41	0,91	-2,14 – 3,12
NSB1, porsas	309	-0,01	0,41	-0,93 – 1,33
NSB2, porsas	309	0,00	0,38	-0,89 – 1,15
PM1, porsas	309	-0,09	0,45	-2,47 – 1,47
PM2, porsas	309	0,00	0,48	-1,89 – 1,71
AFF, vrk	304	20,61	12,64	-25,48 – 60,87
FFI, vrk	270	5,03	5,57	-10,55 – 23,15
SFI, vrk	268	2,45	5,08	-12,89 – 19,32

TNB1 = syntyneiden porsaiden lukumäärä ensimmäisessä pahnueessa, TNB2 = syntyneiden porsaiden lukumäärä myöhemmissä pahnueissa, NSB1 = kuolleena syntyneiden porsaiden lukumäärä ensimmäisessä pahnueessa, NSB2 = kuolleena syntyneiden porsaiden lukumäärä myöhemmissä pahnueissa, PM1 = porsaskuolleisuus syntymästä vieroitukseen ensimmäisessä pahnueessa, PM2 = porsaskuolleisuus syntymästä vieroitukseen myöhemmissä pahnueissa, AFF = emakon ikä ensimmäisellä porsimisella, FFI = ensimmäinen porsimisväli, SFI = toinen porsimisväli

N = havaintojen lukumäärä, μ = keskiarvo, σ = keskihajonta,

Min. – max. = pienin arvo – suurin arvo

5.2 Assosiaatioanalyysi

Tilastollisesti merkitsevät SNP-markkerit olivat yhteydessä ensimmäisessä pahnueessa syntyneiden porsaiden lukumäärään (TNB1), myöhemmissä pahnueissa syntyneiden porsaiden lukumäärään (TNB2), ensimmäisessä pahnueessa kuolleena syntyneiden porsaiden lukumäärään (NSB1), ensimmäisen pahnueen syntymän ja vieroituksen väliseen porsaskuolleisuuteen (PM1), emakon ikään ensimmäisellä porsimisella (AFF) ja toiseen porsimisväliin (SFI) (Taulukko 3). Suurin osa näistä markkereista sijaitsi kromosomissa 7. Myös kromosomeista 13, 8 ja 1 löydettiin tilastollisesti merkitseviä SNP-markkereita. PM1-,

TNB2-, AFF- ja SFI-ominaisuuksiin liittyi kuhunkin useampi merkitsevä SNP-markkeri. PM1:een liittyvät kolme merkitsevää SNP-markkeria sijaitsivat kaikki kromosomissa 13, 90 Mb:n alueella. AFF:hin liittyvät SNP:t sijaitsivat 20 – 30 Mb:n alueella kromosomissa 7 (Taulukko 3). Positiiviset regressiokertoimet (b) taulukossa 3 tarkoittavat sen verran enemmän porsaita (TNB1, TNB2, NSB1, NSB2 ja PM1) tai päiviä (AFF ja SFI).

Taulukko 3. Hedelmällisyysominaisuuksien tilastollisesti merkitsevien ja suuntaa antavien SNP:iden alleelivaikutukset (b) ja P-arvot. Tilastollisesti merkitsevä P-arvo $\leq 2,0E-06$ (lihavoidut P-arvot), suuntaa-antava tulos (P-arvo $\leq 4,0E-06$).

Omin.	SNP	Krom.	Sijainti (bp)	CR	MAF	N	b	SE	P-arvo
TNB1	H3GA0021970	7	66804136	1,0	0,33	310	-0,43	0,06	8,39E-08
TNB2	H3GA0021970	7	66804136	1,0	0,33	310	-0,42	0,06	2,77E-07
TNB2	CASI0009733	1	106461295	0,98	0,12	303	-0,57	0,09	1,91E-06
TNB2	MARC0028113	1	102037572	1,0	0,11	310	-0,57	0,10	3,46E-06
NSB1	H3GA0025264	8	89641384	1,0	0,48	309	-0,17	0,02	9,72E-08
NSB1	ALGA0049016	8	90815538	1,0	0,44	310	-0,15	0,02	2,21E-06
NSB1	ALGA0049057	8	91366284	1,0	0,44	310	-0,15	0,02	2,21E-06
NSB2	ASGA0039145	8	70693590	0,99	0,14	308	-0,20	0,03	3,03E-06
PM1	ALGA0071853	13	92718203	0,98	0,14	304	-0,26	0,03	2,12E-07
PM1	ALGA0071884	13	93747975	1,0	0,13	310	-0,25	0,03	7,01E-07
PM1	ASGA0058763	13	93065294	1,0	0,13	310	-0,25	0,03	7,01E-07
AFF	ALGA0039480	7	23743424	1,0	0,23	310	6,78	0,86	3,89E-08
AFF	ALGA0039780	7	29445530	0,99	0,25	307	6,56	0,85	5,49E-08
AFF	M1GA0009779	7	27855864	1,0	0,28	310	6,17	0,83	1,23E-07
AFF	ASGA0032033	7	27837029	1,0	0,28	309	6,14	0,83	1,42E-07
AFF	ALGA0039427	7	22860527	1,0	0,25	310	6,27	0,84	1,87E-07

AFF	ALGA0039615	7	27133720	0,98	0,35	305	5,68	0,78	2,30E-07
AFF	H3GA0020436	7	27565886	1,0	0,26	309	6,00	0,84	3,36E-07
AFF	H3GA0020344	7	24517046	1,0	0,26	310	5,96	0,84	3,91E-07
AFF	ALGA0039782	7	29466918	0,99	0,32	307	-5,50	0,78	8,86E-07
AFF	MARC0055565	7	24629488	1,0	0,33	310	-5,42	0,80	1,43E-06
AFF	ASGA0031930	7	24853094	1,0	0,27	310	5,57	0,84	1,74E-06
SFI	INRA0028855	7	128927135	0,99	0,06	307	4,16	0,37	4,03E-07
SFI	H3GA0023563	7	129032719	1,0	0,06	310	4,05	0,36	4,03E-07

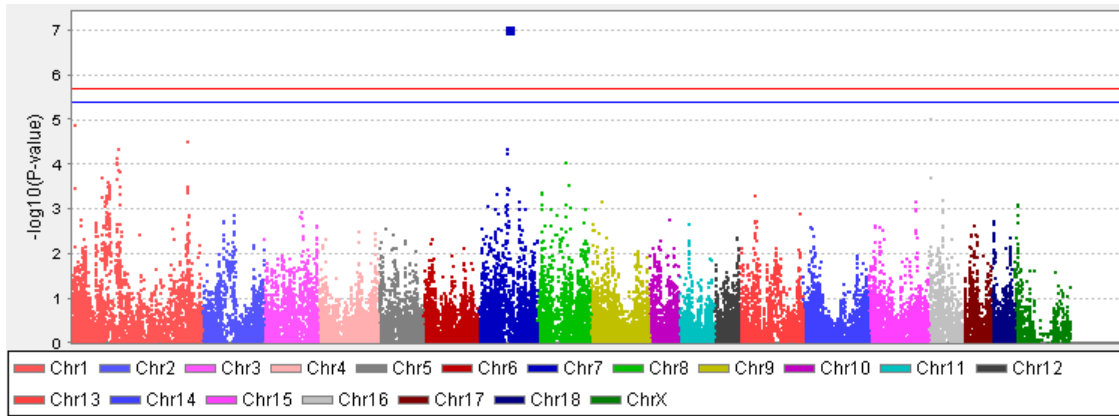
TNB1 = syntyneiden porsaiden lukumäärä ensimmäisessä pahnueessa, TNB2 = syntyneiden porsaiden lukumäärä myöhemmissä pahnueissa, NSB1 = kuolleena syntyneiden porsaiden lukumäärä ensimmäisessä pahnueessa, NSB2 = kuolleena syntyneiden porsaiden lukumäärä myöhemmissä pahnueissa, PM1 = porsaskuolleisuus syntymästä vieroitukseen ensimmäisessä pahnueessa, PM2 = porsaskuolleisuus syntymästä vieroitukseen myöhemmissä pahnueissa, AFF = emakon ikä ensimmäisellä porsimisella, FFI = ensimmäinen porsimisväli, SFI = toinen porsimisväli

CR = call rate-arvo, MAF = populaatiossa vähemmän esiintyvän alleelin frekvenssi, N = havaintojen lukumäärä, b = regressiokerroin, SE = keskivirhe

Seuraavissa kuvissa (Kuvat 1 – 9) esiintyvä punainen linja on tilastollisesti merkitsevän SNP-markkerin raja (P-arvo = 2,0E-06) ja sininen linja on suuntaa antavan SNP-markkerin raja (P-arvo = 4,0E-06).

5.2.1 Ensimmäisessä pahnueessa syntyneiden porsaiden lukumäärä (TNB1)

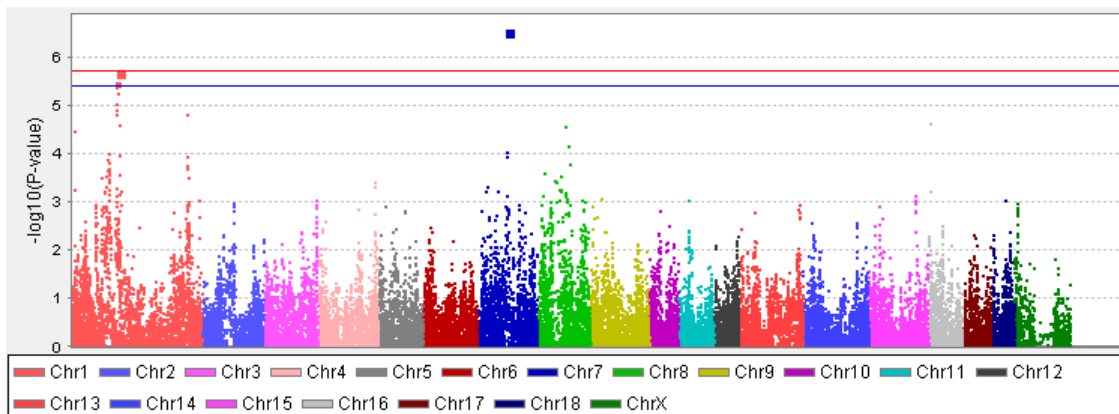
TNB1-ominaisuuteen liittyi yksi (H3GA0021970) tilastollisesti merkitsevä (P-arvo = 8,39E-08) SNP-markkeri (Kuva 1 ja Taulukko 3). Tilastollisesti merkitsevien SNP-markkerien (Taulukko 3) havaittu harvinaisemman alleelin vaikutus oli negatiivinen (0,43 porsasta vähemmän).



Kuva 1. P-arvojen ($-\log_{10}$) jakauma kromosomeittain ensimmäisessä pahnueessa syntyneiden porsaiden lukumäärän (TNB1) suhteen.

5.2.2 Myöhemmissä pahnueissa syntyneiden porsaiden lukumäärä (TNB2)

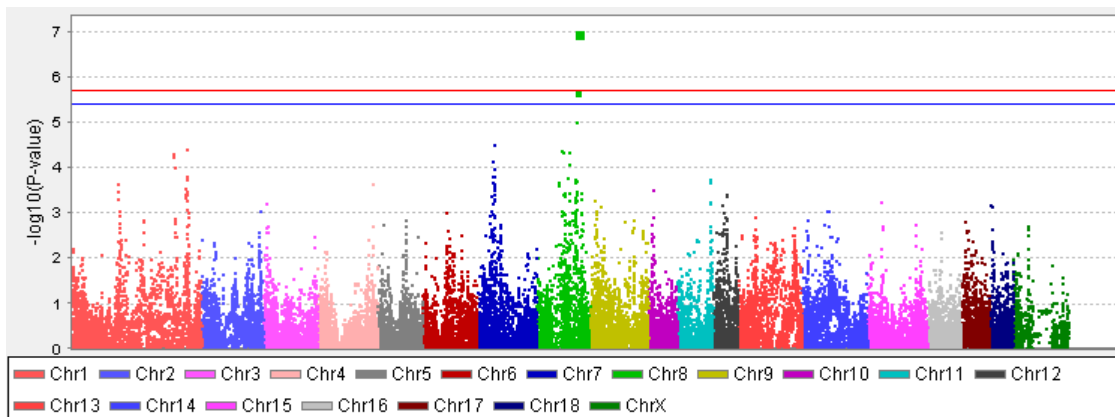
TNB2-ominaisuuteen löytyi kaksi (H3GA0021970 ja CASI0009733) merkitsevää (P-arvot = $2,77E-07$ ja $1,91E-06$) SNP:tä ja yksi suuntaa antava (P-arvo = $3,46E-06$) SNP (Kuva 2 ja Taulukko 3). TNB2:n SNP:iden vaikutus vaihteli $-0,42$ ja $-0,57$ porsaan välillä eli harvinaisemman alleelin vaikutus oli pahnuekokoa pienentävä.



Kuva 2. P-arvojen ($-\log_{10}$) jakauma kromosomeittain myöhemmissä pahnueissa syntyneiden porsaiden lukumäärän (TNB2) suhteen.

5.2.3 Ensimmäisessä pahnueessa kuolleina syntyneiden porsaiden lukumäärä (NSB1)

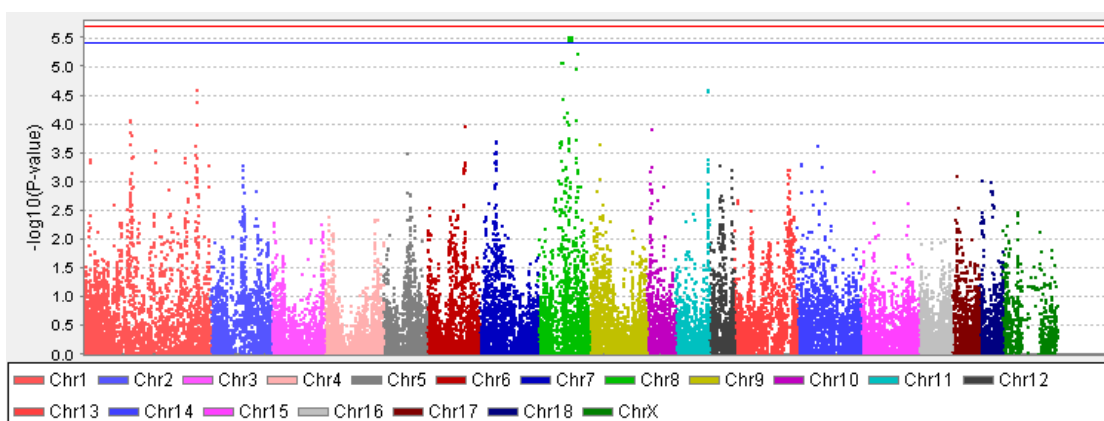
NSB1:een liittyi yksi (H3GA0025264) merkitsevä (P-arvo = $9,72E-08$) ja kaksi suuntaa antavaa (kummallakin P-arvo = $2,21E-06$) SNP:iä (Kuva 3 ja Taulukko 3). Havaittu harvinaisemman alleelin vaikutus oli 0,17 kuollutta porsasta vähemmän.



Kuva 3. P-arvojen ($-\log_{10}$) jakauma kromosomeittain ensimmäisessä pahnueessa kuolleina syntyneiden porsaiden lukumäärän (NSB1) suhteen.

5.2.4 Myöhemmissä pahnueissa kuolleina syntyneiden porsaiden lukumäärä (NSB2)

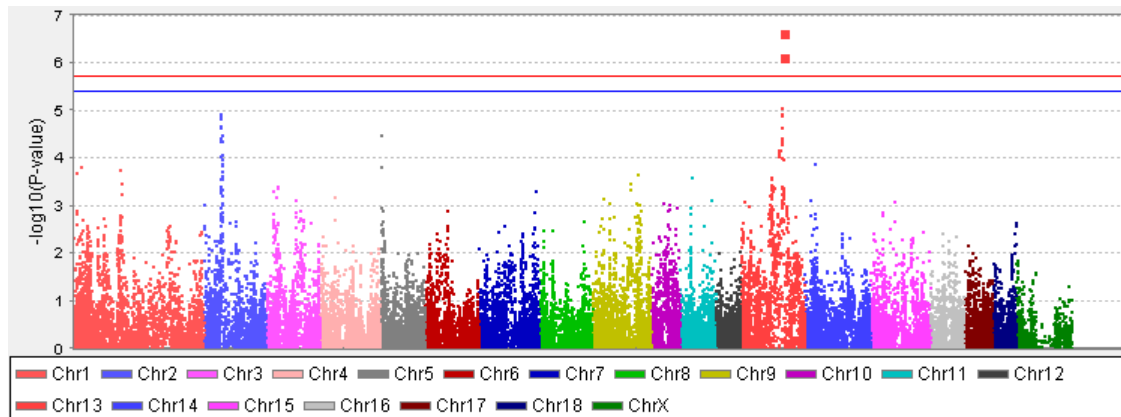
NSB2:een liittyi vain yksi suuntaa antava (P-arvo = $3,03E-06$) SNP-markkeri. (Kuva 4 ja Taulukko 3).



Kuva 4. P-arvojen ($-\log_{10}$) jakauma kromosomeittain myöhemmissä pahnueissa kuolleina syntyneiden porsaiden lukumäärän (NSB2) suhteen.

5.2.5 Porsaskuolleisuus syntymästä vieroitukseen ensimmäisessä pahnueessa (PM1)

PM1:een liittyi kolme (ALGA0071853, ALGA0071884 ja ASGA0058763) tilastollisesti merkitsevää (P-arvo = $2,12E-07$ ja kahdella jälkimmäisellä molemmilla $7,01E-07$) SNP-markkeria (Kuva 5 ja Taulukko 3).

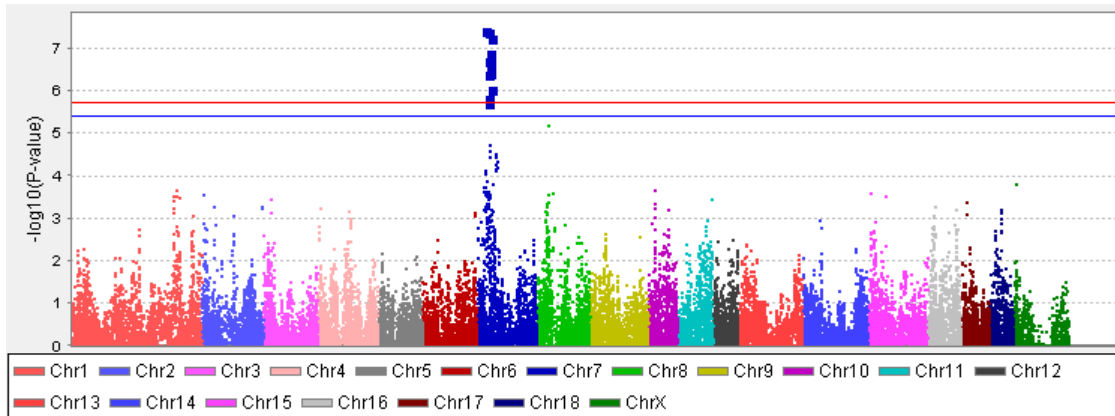


Kuva 5. P-arvojen ($-\log_{10}$) jakauma kromosomeittain ensimmäisen pahnueen syntymän ja vieroituksen välisen porsaskuolleisuuden (PM1) suhteen.

Kolme merkitsevää SNP-markkeria vaikuttivat porsaskuolleisuuteen (PM1), niin että porsaita kuolee 0,25 - 0,26 vähemmän. Nämä SNP:t sijaitsevat kromosomissa 13, 93-94 Mb:n alueella.

5.2.6 Emakon ikä ensimmäisellä porsimisella (AFF)

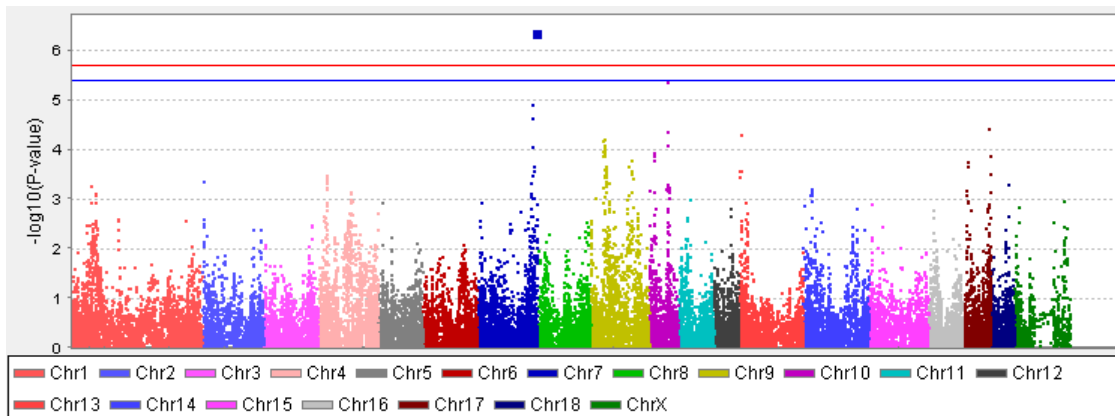
AFF:een liittyi 11 tilastollisesti merkitsevää (p-arvot $1,74E-06$ ja $3,89E-08$ välillä) SNP-markkeria (Kuva 6 ja Taulukko 3). Näiden SNP:iden alleelivaikutus vaihteli -5,50 ja 6,78 päivän välillä.



Kuva 6. P-arvojen ($-\log_{10}$) jakauma kromosomeittain emakon ikä ensimmäisellä porsimisella-ominaisuuden (AFF) suhteen.

5.2.7 Toinen porsimisväli (SFI)

SFI:een liittyi kaksi (INRA0028855 ja H3GA0023563) merkitsevää (molemmilla P-arvot $4,03E-07$) SNP:tä (Kuva 7 ja Taulukko 3). Näiden SNP:iden alleelivaikutus oli 4,05 – 4,16 päivää.



Kuva 7. P-arvojen ($-\log_{10}$) jakauma kromosomeittain toisen porsimisvälin (SFI) suhteen.

6 TULOSTEN TARKASTELU

Vaikka tutkimuksen aineisto oli pieni (310 karjun tiedot), saadut tulokset osoittavat, että käytettyjen karjujen määrä oli riittävä sian hedelmällisyysominaisuuksiin vaikuttavien kromosomialueiden löytämiseksi. Kromosomeista 7, 1, 8 ja 13 löydettiin uusia alueita, jotka vaikuttavat pahnuekokoon, porsaskuolleisuuteen ja emakon porsimisikään. Jotta

pahnuekokoon, porsaskuolleisuuteen ja emakon porsimisikään liittyvien kromosomialueiden geenivaikutukset pystytään arvioimaan tarkemmin, tarvitaan laajempia populaatiotutkimuksia.

Tutkimuksen SNP-markkereiden sekvenssikuvaus perustui sian genomien koontiin 9. Koonti 9 sisältää vielä monia virheitä liittyen SNP:iden tarkkaan sijaintiin ja niiden järjestykseen. Tutkimuksen mukaiset SNP-assosiaatiotestit eivät ole riippuvaisia käytetystä koonnista. Tutkimuksessa SNP:ille saadut P-arvot pitävät paikkansa, vaikka käytettäisiin uudempaa koontia. Tilastollisesti merkitsevien SNP:iden läheisyydessä olevat geenit saattavat kuitenkin muuttua käytettäessä uudempaa koontia.

6.1 Kirjallisuudessa raportoidut QTL:t

Kromosomista 7 on raportoitu kaksi QTL:ää (26 cM (Li ym. 2009) ja 32 cM (Buske ym. 2005)), jotka ovat yhteydessä syntyneiden porsaiden lukumäärään. Tässä tutkimuksessa löydetty, TNB1-ominaisuuteen yhteydessä oleva SNP sijaitsee kromosomissa 7, 67 Mb:n kohdalla. Nyt löydetty alue on uusi eikä sitä ole raportoitu aikaisemmin. Myös kromosomista 1 on raportoitu kaksi QTL:ää (kummankin sijainti 19 cM (Rothschild ym. 1996 ja Horogh ym. 2005)), jotka ovat yhteydessä syntyneiden porsaiden lukumäärään. Tämän tutkimuksen TNB2-ominaisuuteen vaikuttava tilastollisesti merkitsevä SNP sijaitseekin kromosomissa 1 (106 Mb) se ei siis ole sama aikaisemmin raportoitujen löydösten kanssa.

Kromosomista 8 on raportoitu yksi kuolleina syntyneiden porsaiden lukumäärään liittyvä QTL (73 cM (Li ym. 2009)). Tämän tutkimuksen NSB1-ominaisuuden tilastollisesti merkitsevä SNP sijaitsee 90 Mb:n alueella, joten havaittu ja raportoitu alue voivat olla samoja ottaen huomioon, että raportoitu QTL on löydetty kytkentäanalyyysin avulla (risteytysaineisto), jossa luottamusvälin suuruus on yleensä kymmeniä cM:iä eli kymmeniä miljoonia emäspareja (Mb).

Kromosomista 13 ei ole raportoitu kuolleina syntyneiden porsaiden lukumäärään (PM) liittyviä QTL:iä.

Kromosomista 7 on raportoitu viisi emakon puberteetti-ikään yhteydessä olevaa QTL:ää (sijainnit: 5 cM ja 45 cM (Cassady ym. 2001), 49 cM (Holl ym. 2004), 50 cM (Bidanel ym.

2008) ja 59 cM (Yang ym. 2008)). Tämän tutkimuksen, emakon porsimisikään (AFF) liittyvät SNP:t sijaitsevat kromosomissa 7, 23 ja 30 Mb:n välisellä alueella. Havaittu ja osa raportoiduista alueista voivat olla samoja, sillä raportoidut QTL:t (Cassady ym. 2001, Holl ym. 2004 ja Bidanel ym. 2008) on myös löydetty kytkentäanalyysin avulla eli niihin liittyvä luottamusväli on suuri.

Kromosomista 7 ei ole raportoitu porsimisväliin (FFI tai SFI) liittyviä QTL:itä.

6.2 Mahdollinen kandidaattigeeni

NSB1 Kromosomissa 8 sijaitseva alue (sijainti: 89,6 Mb, H3GA0025264-SNP:n läheisyydessä) sisältää yhden potentiaalisen kandidaattigeenin (MAD2L1), jolla saattaa olla vaikutusta emakon hedelmällisyyteen. Wangin ja Sunin (2006) mukaan viimeaikaiset tutkimukset osoittavat, että sukkularihmaston tarkastuspiste-järjestelmä osallistuu nisäkäsnaaraiden meioosin säätelyyn ja alkioasteella ilmenevän poikkeavan kromosomiluvun syntymisen ehkäisyyn. Mitoosin aikana sukkularihmaston tarkastuspiste-proteiini MAD2 ehkäisee aneuploidiaa eli poikkeavaa kromosomiluvun syntymistä (Homer ym. 2005). Ihmisillä meioosissa 1 tapahtuvat virheet ovat keskenmenojen ja henkisen jälkeenjääneisyyden tärkein geneettinen aiheuttaja (Hassold ja Hunt 2001).

Tutkimuksessa ei löytynyt sian hedelmällisyysominaisuuksille muita mahdollisia kandidaattigeenejä.

6.3 Vertailu maatiaisrodun tuloksiin

Maatiaisrodun kaikki tilastollisesti merkitsevät alueet sijaitsivat kromosomissa 9 (Taulukko 4). Tämä tarkoittaa sitä, että Suomen yorkshire-rodun ja maatiaisrodun hedelmällisyysominaisuuksiin eivät vaikuta samat kromosomialueet.

Taulukko 4. Maatiaisrodun tilastollisesti merkitsevät SNP:t (Uimari ym. 2011)

Tilastollisesti merkitsevä $\leq 2,0E-06$, TNB1 = syntyneiden porsaiden lukumäärä ensimmäisessä pahnueessa, TNB2 = syntyneiden porsaiden lukumäärä myöhemmissä pahnueissa, PM2 = porsaskuolleisuus syntymästä vieroitukseen myöhemmissä pahnueissa

Ominaisuus	Merkitsevä SNP	P-arvo	Kromosomi	Sijainti (Mb)
TNB1	DRGA0009645	3,21E-07	9	95
TNB1	ALGA0054078	4,24E-07	9	79
TNB1	H3GA0027863	8,21E-07	9	79
TNB1	MARC0003458	8,21E-07	9	79
TNB1	MARC0027588	8,21E-07	9	79
TNB2	DRGA0009645	8,45E-08	9	95
TNB2	ALGA0054078	5,62E-07	9	79
TNB2	H3GA0027863	1,03E-06	9	79
TNB2	MARC0003458	1,03E-06	9	79
TNB2	MARC0027588	1,03E-06	9	79
PM2	ASGA0043706	6,94E-08	9	65
PM2	MARC0027886	7,98E-08	9	66
PM2	ALGA0053783	8,85E-07	9	66

6.4. Tulosten hyödyntäminen käytännössä

Tutkimuksen tuloksia voidaan hyödyntää merkkiavusteisessa valinnassa. Tällöin sikoja voidaan valita hedelmällisyysominaisuuksien kannalta suotuisien SNP-markkereiden suhteen, jolloin hedelmällisyysominaisuuksien geneettinen edistyminen nopeutuu. Saatujen tulosten avulla voidaan myös selvittää, mitkä geenit vaikuttavat hedelmällisyysominaisuuksiin ja millaiset niiden geenivaikutukset ovat. Tämä edellyttää riittävää kytkentäepätasapainoa (LD) markkerien ja QTL:n välillä.

7 JOHTOPÄÄTÖKSET

Tutkimuksessa löydettiin sian hedelmällisyysominaisuuksiin vaikuttavia tilastollisesti merkitseviä kromosomialueita. Pahnuekokoon, porsaskuolleisuuteen ja emakon porsimisikään vaikuttavat alueet sijaitsevat kromosomeissa 7, 1, 8 sekä 13. Tutkimuksessa löydettiin myös suuntaa antavia alueita kromosomeista 1 ja 8. Nämä kromosomialueet vaikuttavat myöhempien pahnueiden pahnuekokoon ja kuolleena syntyneiden porsaiden lukumäärään ensimmäisessä pahnueessa. Löydettyjen kromosomialueiden alleelivaikutukset ovat kuitenkin melko vähäisiä.

Vertailtaessa Suomen yorkshire-rodun ja maatiaisrodun tuloksia keskenään, saatiin selville, että näiden rotujen hedelmällisyysominaisuuksiin eniten vaikuttavat geenit (kromosomialueet) eivät ole samoja.

Jos tässä tutkimuksessa löydetty kromosomialueet saadaan varmistettua, voidaan näitä SNP-markkereita hyödyntää merkkiavusteisen valinnan avulla kansallisessa jalostusohjelmassa. MAS:n avulla sikoja voidaan alkaa valita hedelmällisyysominaisuuksien kannalta suotuisien SNP-markkereiden suhteen ja pidemmällä tähtäimellä hedelmällisyysominaisuuksien geneettinen edistyminen alkaa nopeutua. Laajempien populaatiotutkimusten avulla voidaan myös selvittää, mitkä geenit vaikuttavat hedelmällisyysominaisuuksiin ja millaiset niiden geenivaikutukset ovat.

8 KIITOKSET

Tutkimus oli osa MTT:n Hyvinvointia ruuasta –ohjelman SikaSNiP-tutkimushanketta. Kiitokset kaikille hankkeessa mukana olleille ja kiitos myös osk. Faballe tutkimuksen aineistosta. Erityiskiitokset ohjaajilleni erikoistutkija Pekka Uimarille ja professori Jarmo Jugalle. Haluan kiittää perhettäni teknisestä tuesta ja erityisesti oikoluvusta. Lämpimät kiitokset ystäväilleni, jotka tukivat minua prosessin aikana.

9 LÄHTEET

- Adamec, V. & Johnson, R. K. 1997. Genetic analysis of rebreeding intervals, litter traits, and production traits in sows of the National Czech nucleus. *Livestock Production Science* 48: 13-22.
- Alfonso, L., Noguera, J. L., Babot, D. & Estany, J. 1997. Estimates of genetic parameters for litter size at different parities in pigs. *Livestock Production Science* 47: 149-156.
- Amaral, A. J., Megens, H-J., Crooijmans, R. P. M. A., Heuven, H. C. M. & Groenen, M. A. M. 2008. Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics Society of America* 179: 569-579.
- Amaral, A. J., Megens, H-J., Kerstens, H. H. D., Heuven, H. C. M., Dibbits, B., Crooijmans, R. P. M. A., den Dunnen, J. T. & Groenen, M. A. M. 2009. Application of massive parallel sequencing to whole genome SNP discovery in the porcine genome. *BioMed Central Genomics* 10: 374.
- Andersson, L. 2001. Genetic dissection of phenotypic diversity in farm animals. *Nature Reviews Genetics* 2: 130-138.
- Andersson, L. 2008. Genome-wide association analysis in domestic animals: a powerful approach for genetic dissection of trait loci. *Genetica* 136: 341-349.
- Archibald, A. L., Bolund, L., Churcher, C., Fredholm, M., Groenen, M. A. M., Harlizius, B., Lee, K-T., Milan, D., Rogers, J., Rothschild, M. F., Uenishi, H., Wang, J., Schook, L. B. & the Swine Genome Sequencing Consortium. 2010. *BioMed Central Genomics* 11: 438.
- Balding, D. J. 2006. A tutorial on statistical methods for population association studies. *Nature Reviews Genetics* 7: 781-791.
- Bennett, G. L. & Leymaster, K. A. 1989. Integration of ovulation rate, potential embryonic viability and uterine capacity into a model of litter size in swine. *Journal of Animal Science* 67: 1 230-1 241.
- Bidanel, J. P., Rosendo, A., Iannuccelli, N., Riquet, J., Gilbert, H., Caritez, J. C., Billon, Y., Amigues, Y., Prunier, A. & Milan, D. 2008. Detection of quantitative trait loci for teat number and female reproductive traits in Meishan x Large White F2 pigs. *Animal* 2: 813-820.
- Buske, B., Brunsch, C., Zeller, K., Reinecke, P. & Brockmann, G. 2005. Analysis of properdin (*BF*) genotypes associated with litter size in a commercial pig cross population. *Journal of Animal Breeding and Genetics* 122: 259-263.
- Calus, M. P. L., Meuwissen, T. H. E., de Roos, A. P. W. & Veerkamp, R. F. 2008. Accuracy of genomic selection using different methods to define haplotypes. *Genetics* 178: 553–561.

- Cassady, J. P., Johnson, R. K., Pomp, D., Rohrer, G. A., Van Vleck, L. D., Spiegel, E. K. & Gilson, K. M. 2001. Identification of quantitative trait loci affecting reproduction in pigs. *Journal of Animal Science* 79: 623–633.
- Chen, P., Baas, T. J., Mabry, J. W., Koehler, K. J. & Dekkers, J. C. M. 2003. Genetic parameters and trends for litter traits in U. S. Yorkshire, Duroc, Hampshire, and Landrace pigs. *Journal of Animal Science* 81: 46–53.
- Collins, A., Lonjou, C. & Morton, N. E. 1999. Genetic epidemiology of single-nucleotide polymorphisms. *Proceedings of the National Academy of Sciences of the United States of America* 96: 15173-15177.
- Dekkers, J. C. M. 2004. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *Journal of Animal Science* 82: 313-328.
- Drogemuller, C., Hamann, H. & Distl, O. 2001. Candidate gene markers for litter size in different German pig lines. *Journal of Animal Science* 79: 2 565-2 570.
- Du, F-X., Clutter, A. C. & Lohuis, M. M. 2007. Characterizing linkage disequilibrium in the pig populations. *International Journal of Biological Sciences* 3: 166-178.
- Ensembl 2009. *Sus scrofa*. http://www.ensembl.org/Sus_scrofa/Info/Index. Päivitetty toukokuussa 2012, julkaistu 2009, viitattu 19.10.2011.
- Ensembl 2009. *Sus scrofa*, archives. http://www.ensembl.org/Sus_scrofa/Info/StatsTable?db=core. Päivitetty helmikuussa 2012, viitattu 21.2.2012.
- FABA Palvelu 2010. Toimintakertomus ja tilastot 2009. http://www.faba.fi/files/1702/Faba_Palvelun_toimintakertomus_2009.pdf. Julkaistu 2010, viitattu 20.2.2012.
- Fries, R. & Durstewitz, G. 2001. Digital DNA signatures: SNPs for animal tagging. *Nature Biotechnology* 19: 508.
- Garrick, D. J. 2011. Review. The nature, scope and impact of genomic prediction in beef cattle in the United States. *Genetics Selection Evolution* 43: 17.
- Garrick, D. J., Taylor, J. F. & Fernando, R. L. 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genetics Selection Evolution* 41: 55.
- Goddard, M. E. & Hayes, B. J. 2007. Genomic selection. *Journal of Animal Breeding and Genetics* 124: 323-330.
- Haltia, S. 2010. Pahnuekoot kasvussa. *Sika* 3 (2010): 10-12.
- Hassold, T. & Hunt, P. 2001. To err (meiotically) is human: The genesis of human aneuploidy. *Nature Reviews Genetics* 2: 280-291.

- Heaton, M. P., Harhay, G. P., Bennett, G. L., Stone, R. T., Grosse, W. M., Casas, E., Keele, J. W., Smith, T. P. L., Chitko-McKown, C. G. & Laegreid, W. W. 2002. Selection and use of SNP markers for animal identification and paternity analysis in U.S. beef cattle. *Mammalian Genome* 13: 272-281.
- Holl, J. W., Cassady, J. P., Pomp, D. & Johnson, R. K. 2004. A genome scan for quantitative trait loci and imprinted regions affecting reproduction in pigs. *Journal of Animal Science* 82: 3421-3429.
- Homer, H. A., McDougall, A., Levasseur, M., Yallop, K., Murdoch, A. P. & Herbert, M. 2005. Mad2 prevents aneuploidy and premature proteolysis of cyclin B and securin during meiosis 1 in mouse oocytes. *Genes & Development* 19: 202-207.
- Horogh, G., Zsolnai, A., Komlósi, I., Nyíri, A., Anton, I. & Fésüs, L. 2005. Oestrogen receptor genotypes and litter size in Hungarian Large White pigs. *Journal of Animal Breeding and Genetics* 122: 56-61
- Jiang, Z., Gibson, J. P., Archibald, A. L. & Haley, C. S. 2001. The porcine gonadotropin-releasing hormone receptor gene (GNRHR): Genomic organization, polymorphisms, and association with the number of corpora lutea. *Genome* 44: 7-12.
- Jungerius, B. J., Gu, J., Crooijmans, R. P. M. A., van der Poel, J. J., Groenen, A. M., van Oost, B. A. & te Pas, M. F. W. 2005. Estimation of the extent of linkage disequilibrium in seven regions of the porcine genome. *Animal Biotechnology* 16: 41-54.
- Kadarmideen, H. N., von Rohr, P. & Janss, L. L. G. 2006. From genetical genomics to systems genetics: potential applications in quantitative genomics and animal breeding. *Mammalian Genome* 17: 548-564.
- King, A. H., Jiang, Z., Gibson, J. P. Haley, C. S. & Archibald, A. L. 2003. Mapping quantitative trait loci affecting female reproductive traits on porcine chromosome 8. *Biology of Reproduction* 68: 2 172-2 179.
- Lamberson, W. R., Johnson, R. K., Zimmerman, D. R. & Long, T. E. 1991. Direct responses to selection for increased litter size, decreased age at puberty, or random selection following selection for ovulation rate in swine. *Journal of Animal Science* 69: 3 129-3 143.
- Lande, R. & Thompson, R. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124: 743-756.
- Li, X. P., Hu, Z. L., Moon, S. J., Do, K. T., Ha, Y. K., Kim, H., Byun, M. J., Choi, B. H., Rothschild, M. F., Reecy, J. M. & Kim, K. S. 2008. Development of an in silico coding gene SNP maps in pigs. *Animal Genetics* 39: 446-450.
- Li, K., Ren, J., Xing, Y., Zhang, Z., Ma, J., Guo, Y. & Huang, L. 2009. Quantitative trait loci for litter size and prenatal loss in a White Duroc × Chinese Erhualian resource population. *Animal Genetics* 40: 963-966.
- Lund, M. S., Puonti, M., Rydhmer, L. & Jensen, J. 2002. Relationship between litter size and perinatal and pre-weaning survival in pigs. *Animal Science* 74: 217-222.

- Madsen, P., Su, G., Labouriau, R. & Christensen, O. F. 2006. DMU - A package for analyzing multivariate mixed models. 8th World Congress on Genetics Applied to Livestock Production.
- Meuwissen, T. H. E., Hayes, B. & Goddard, M. E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819 - 1829.
- NAGP Bioinformatics team 2003. PigQTLdb. <http://www.genome.iastate.edu/cgi-bin/QTLdb/SS/index>. Päivitetty huhtikuussa 2012, viitattu 28.3.2012.
- National Center for Biotechnology Information 1993. NSBL. <http://www.ncbi.nlm.nih.gov/pubmed/> Päivitetty helmikuussa 2012, viitattu 17.8.2012.
- Ostersen, T., Christensen, O. F., Henryon, M., Nielsen, B., Su, G. & Madsen, P. 2011. Deregressed EBV as the response variable yield more reliable genomic predictions than traditional EBV in pure-bred pigs. *Genetics Selection Evolution* 43: 38.
- Patterson, H. D. & Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* 58: 545-554.
- Ramos, A. M., Crooijmans, R. P. M. A., Affara, N. A., Amaral, A. J., Archibald, A. L., Beever, J. E., Bendixen, C., Churcher, C., Clark, R., Dehais, P., Hansen, M. S., Hedegaard, J., Hu, Z-L., Kerstens, H. H., Law, A. S., Megens, H-J., Milan, D., Nonneman, D. J., Rohrer, G. A., Rothschild, M. F., Smith, T. P. L., Schnabel, R. D., Van Tassel, C. P., Taylor, J. F., Wiedmann, R. T., Schook, L. B. & Groenen, M. A. M. 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS ONE* 4.
- Rathje, T. A., Rohrer, G. A. & Johnson, R. K. 1997. Evidence for quantitative loci affecting ovulation rate in pigs. *Journal of Animal Science* 75: 1 486–1 494.
- Rempel, L. A., Nonneman, D. J., Wise, T. H., Erkens, T., Peelman, L. J. & Rohrer, G. A. 2010. Association analyses of candidate single nucleotide polymorphisms on reproductive traits in swine. *Journal of Animal Science* 88: 1-15.
- Rohrer, G. A. Ford, J. J. Wise, T. H. Vallet, J. L. & Christenson, R. K. 1999. Identification of quantitative trait loci affecting female reproductive traits in a multigeneration Meishan-White composite swine population. *Journal of Animal Science* 77: 1 385–1 391.
- Rothschild, M. Jacobson, C., Vaske, D., Tuggle, C., Wang, L., Shorts, T., Eckardt, G., Sasaki, S., Vincent, A., McLaren, D., Southwood, O., Van der Steen, H., Mileham, A. & Plastow, G. 1996. The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proceedings of the National Academy of Sciences of the United States of America* 93: 201-205.
- Rothschild, M. F. & Soller, M. 1997. Candidate gene analysis to detect genes controlling traits of economic importance in domestic livestock. *Probe* 8: 13-20.
- Schaeffer, L. R. 2006. Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics* 123: 218-223.

- Schook, L. B., Beever, J. E., Rogers, J., Humphray, S., Archibald, A., Chardon, P., Milan, D., Rohrer, G. & Eversole, K. 2005. Swine genome sequencing consortium (SGSC): a strategic roadmap for sequencing the pig genome. *Comparative and Functional Genomics* 6: 251-255.
- Schulman, N. F., Sahana, G., Iso-Touru, T., McKay, S. D., Schnabel, R. D., Lund, M. S., Taylor, J. F., Virta, J. & Vilkki, J. H. 2011. Mapping of fertility traits in Finnish Ayrshire by genome-wide association analysis. *Animal Genetics* 42: 263–269.
- Serenius, T. 2004. Genetics of sow efficiency in the Finnish Landrace and Large White populations. *Agrifood Research Reports* 55: 1-41.
- Simianer, H. & Pimentel, E. C. G. 2009. Robust QTL fine mapping by applying a quantitative transmission disequilibrium test to the Mendelian sampling term. *Journal of Animal Breeding and Genetics* 126: 432-442.
- Strandén, I. & Garrick, D. J. 2009. Technical note: Derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. *Journal of Dairy Science* 92: 2 971–2 975.
- Strauss, S. 2010. Biotech breeding goes bovine. *Nature Biotechnology* 28: 540-543.
- Uimari, P., Sironen, A. & Sevón-Aimonen, M-L. 2011. Whole-genome SNP association analysis of reproduction traits in the Finnish Landrace pig breed. *Genetics Selection Evolution* 43: 42-49.
- Uimari, P. & Tapio, M. 2011. Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. *Journal of Animal Science* 89: 609-614.
- Voight, B. F. & Pritchard, J. K. 2005. Confounding from cryptic relatedness in case-control association studies. *PLoS Genetics* 1: 302–311.
- Wang, W-H. & Sun, Q-Y. 2006. Meiotic spindle, spindle checkpoint and embryonic aneuploidy. *Frontiers in Bioscience* 11: 620-636.
- Wiedmann, R. T., Smith, T. P. L. & Nonneman, D. J. 2008. SNP discovery in swine by reduced representation and high throughput pyrosequencing. *BioMed Central Genomics* 9: 81.
- Wilkie, P. J., Paszek, A. A., Beattie, C. W., Alexander, L. J., Wheeler, M. B. & Schook, L. B. 1999. A genomic scan of porcine reproductive traits reveals possible quantitative trait loci (QTLs) for number of corpora lutea. *Mammalian Genome* 10: 573–578.
- Yang, G., Ren, J., Li, S., Mao, H., Guo, Y., Zou, Z., Ren, D., Ma, J. & Huang, L. 2008. Genome-wide identification of QTL for age at puberty in gilts using a large intercross F2 population between White Duroc x Erhualian. *Genetics Selection Evolution* 40: 529-539.

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S. & Buckler, E. S. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38: 203–208.